EXHIBIT 5

Rebuttal to Reports of Dr. Alex Spiliotopoulos and Dr. Remy J.-C. Hennet

Leonard F. Konikow

January 13, 2025

Qualifications:

I received a PhD in Geosciences from Penn State University in 1973, specializing in hydrogeology and groundwater modeling. I worked as a research hydrologist for the U.S. Geological Survey for about 42 years, and was the Editor-in-Chief of *Groundwater* journal for four years (2020-2023). At the USGS, I was mostly involved in the development, documentation, and application of groundwater flow models and groundwater solute-transport models. I was elected to the National Academy of Engineering in 2015. I am a Fellow of the American Geophysical Union and the Geological Society of America, which also presented me with their Meinzer Award for publications that have significantly advanced the science of hydrogeology. I have served on several Expert Peer Review Panels during my career, including those for ATSDR's Camp Lejeune groundwater modeling studies in 2005 and in 2009.

My curriculum vitae is included with this report as **Attachment A**, and a list of the publications I authored in the previous 10 years is included as **Attachment B**. I am being compensated at an hourly rate of $400 for my work on this litigation. I have not testified at a deposition or trial in the last 4 years.

Introduction:

ATSDR prepared reports describing models developed to simulate groundwater flow and contaminant transport at two areas of Camp Lejeune, North Carolina: Tarawa Terrace (TT) and Hadnot Point/Holcomb Boulevard (HPHB). Their use of the models was innovative in the sense that instead of a typical use of a groundwater model to predict future behavior, they used the model to "predict" how the system evolved in the past (before concentration observations were made) from a known state (an initial condition), in which no contaminants were present, to a contaminated aquifer with a mapped distribution in the early to mid-1980s when contamination was observed at a number of locations (wells, soil samples, and water treatment plants). ATSDR's use of groundwater models to reconstruct trends during a historical gap in concentration measurements is a legitimate and not unprecedented application of groundwater models. In fact, there are other publications in which doing this is documented and considered to be a normal and necessary part of the model calibration process, as discussed in more detail below. Modeling is the best and most logical approach for providing this information.

The ATSDR modeling work was reviewed and commented on by Dr. Alex Spiliotopoulos and Dr. Remy J.C. Hennet. In turn, I was asked to review the reports prepared by Dr. Spiliotopoulos and Dr. Hennet. This report presents my response, comments, and concerns about the technical content of Dr. Spiliotopoulos' and Dr. Hennet's reports. A list of the materials I have considered in rendering my opinions will be provided within seven days.

My opinions expressed in this rebuttal report are based on my review of the reports of Dr. Spiliotopoulos, Dr. Hennet, Mr. Maslia (Oct. 2024), Dr. Aral (Oct. 2024) and Jones & Davis (Oct. 2024), the ATSDR published reports, published literature, documents produced in this litigation, my work on the Camp Lejeune Expert Peer Review Panels, and my experience and expertise in the fields of hydrogeology

and groundwater modeling. I hold these opinions to a reasonable degree of scientific certainty. I reserve the right to supplement and/or amend my opinions in this matter as necessary if additional documents or information are made available for my review.

Background comments about groundwater modeling related to DOJ Expert Reports

This section responds to the opinions of Drs. Spiliotopoulos and Hennet regarding the methodology used by ATSDR to reconstruct groundwater contamination, including their assertions that this methodology is novel, speculative and unfounded, and their repeated claims that this methodology cannot be used where there is limited to no historical data. (*E.g.*, Spiliotopoulos Report, pages 25-30).

A numerical computer model of groundwater flow and/or transport is a simplified representation of a complex reality. A model uses averages, approximations, and assumptions to simulate groundwater behavior and to reproduce its properties and characteristics. Because of uncertainty in defining aquifer properties and boundary conditions, groundwater models must be calibrated. Field observations of aquifer responses (such as changes in water levels for flow models and changes in concentration for transport models) are compared to corresponding model-calculated values. The objective of this calibration procedure is to minimize differences between the observed data and calculated values. The minimization is accomplished by adjusting parameter values within their ranges of uncertainty until a best fit is achieved.

Anderson and Woessner (1992) present a dichotomy of prevailing opinions about mathematical models:
1. "Models are worthless because they require too many data and therefore are too expensive to assemble and run. Furthermore, they can never be proved to be correct and suffer from a lack of scientific certainty."
2. "Models are essential in performing complex analyses and in making informed predictions."
They go on to conclude that "Although groundwater models are time-consuming to design and therefore expensive in terms of labor time, it is also true that use of a groundwater model is the best way to make an informed analysis or prediction about the consequences of a proposed action. … For these reasons, the bias of this book is, of course, toward opinion #2."

Groundwater contamination became widely recognized as a serious and pervasive problem in the 1980s. It is common that the existence of a groundwater contamination problem in a particular area would not be recognized until that contamination has migrated far enough and long enough that it affected a water-supply well or a surface water source. Then a monitoring program might be initiated. But this might not happen for several years to a few decades after the contaminant had entered the aquifer. Therefore, it is common that early-time data on concentrations are simply not available, as is the case at Camp Lejeune. Groundwater modeling is a widely recognized and accepted approach to understanding and managing these contamination problems. Models must be (and have been) calibrated in the absence of early time concentration data, as ATSDR has done. Other representative published examples where this has been successfully accomplished include the Rocky Mountain Arsenal, CO (Konikow, 1977) and Lawrence Livermore National Laboratory, CA (Rogers, 1992). In both of these cases, the early time history was reconstructed as part of the model calibration process (it just wasn't called "hindcasting").

In comparing hindcasting to forecasting, there are some similarities and some differences. In both cases, the analyst is using the model to estimate conditions during a time period outside of the calibration

period, and both types of "predictions" have uncertainty associated with them. One difference is that for predictions of future conditions (forecasting), you can come back later and assess the accuracy of those model predictions. With hindcasting, that is not directly possible. Another difference is that with forecasting (predicting), future conditions are somewhat unbounded, so that uncertainty will tend to increase with time beyond the calibration period. With hindcasting, there is often a way to estimate initial or early time conditions, thereby putting a constraint or bound on uncertainty going back in time. While predictive uncertainty exists and must be recognized, hindcasting is an acceptable and reasonable way to use a calibrated model to assess groundwater conditions during a historical period when there were no observations.

"Hindcasting" was accomplished as part of a study of the Rocky Mountain Arsenal (CO) contamination problem, in which I developed and calibrated a groundwater flow and transport model (Konikow, 1977). The RMA began operations in or about 1943. A groundwater contamination problem was recognized in 1954 & 1955. No observations of concentration (chloride in this early case) were made until late 1955 to 1956. The model was developed to simulate the entire history of operation and contamination at RMA, starting in 1943, but no concentration data were available for the first 13 years of operation. Konikow (1977) made and described reasonable assumptions about the initial conditions, source locations, and source loading—but of course there was uncertainty associated with those estimates (as described by Konikow and Thompson, [1984]). The RMA model was calibrated using measurements made at four distinct times including 1956, 1961, 1969 and 1972. Work was documented and published in a 1977 USGS Water-Supply Paper (https://pubs.usgs.gov/wsp/2044/report.pdf), which received wide distribution. The RMA site became one of the first sites to fall under the Installation Restoration Program. Another example of reconstructing the early history of contamination migration was published by Rogers (1992) in *Groundwater* journal about their model calibration at the Lawrence Livermore National Laboratory site in California. In both of these earlier studies, the historical reconstruction wasn't called "hindcasting," but was considered a scientifically valid component of the model development and application.

Numerical simulation models of groundwater flow and transport processes in porous media are probably the most valuable single tool available to help analysts understand subsurface systems, integrate available data, evaluate conceptual models, and predict responses of groundwater systems to various stresses (such as pumping from wells and leakage or loading of contaminants into the subsurface environment). Groundwater flow models typically estimate the head distribution (equivalent to water levels, water table elevation, or potentiometric surface) in an aquifer system and how the head may change over time in response to changes in well locations or pumping rates. Groundwater transport models (solute transport or contaminant transport for dissolved chemicals) calculate how the concentration of a particular dissolved chemical will vary from place to place and over time. Groundwater systems are three-dimensional in nature, and their properties vary both horizontally and with depth. Therefore, groundwater models must typically be three-dimensional in nature. There is a large record in the published peer-reviewed literature of cases describing the development and application of models for complex real groundwater problems.

Contaminant transport in the subsurface is strongly influenced by the groundwater flow field. Thus, contaminant-transport modeling for a specific site requires a reasonably reliable groundwater flow model. If the contaminant is nonreactive or mildly reactive, the groundwater velocity (based on hydraulic

gradients and effective porosity) is the primary control on advective and dispersive contaminant migration.

Comments about the distribution coefficient (K_d) and the retardation coefficient (R_f)

This section provides background information in support of my responses later in this report to Opinion 3 of Dr. Spiliotopoulos and Opinion 11 of Dr. Hennet regarding the methodology ATSDR used to calculate the retardation factor.

If a contaminant undergoes chemical reactions during the transport process, its net movement relative to the flow of groundwater may be slowed down. Such effects can be (and often are) represented in a simplified manner as a retardation process. Two parameters that are used to simulate retardation are discussed frequently in the comments by Dr. Spiliotopoulos. The contaminant transport conceptual model is that the migration of a contaminant may be slower than the average velocity of the groundwater in which it is dissolved because of adsorption to material in the aquifer. The net effect of this process is described by a so-called "retardation factor" (R_f), which is calculated as:

$$R_f = 1 + (\rho_b\ K_d)/\theta$$

where R_f = retardation factor; ρ_b = bulk density; K_d = distribution coefficient; and θ = porosity.

The model calculates R_f on the basis of the three parameter values on the right side of the above equation, all of which can vary in space and will include uncertainty in their estimated values. If ρ_b is estimated too high by 25% and K_d is too low by 25%, then the errors in those two estimates cancel each other out (i.e., they are compensating errors), and the net estimated value of R_f used in the model will be the same as if those two parameters were estimated precisely to their "true" values.

In general, the use of a distribution coefficient (K_d) as a component of a retardation factor in contaminant transport modeling in groundwater systems is a common modeling approach in simulating contaminant transport in aquifers, but one whose rigorous scientific basis is debatable. The K_d approach assumes that sorption of the PCE is instantaneous, reversible, and follows a linear equilibrium isotherm, and that "the solid matrix has an infinite sorption capacity" (Zhang & Bennett, 2002, p. 81). But in transport through complex heterogeneous porous media, the actual behavior of PCE would not match these idealized assumptions. Nevertheless, it is a simplifying assumption that can be useful in light of the uncertainties about the contaminant's distribution and reactive behavior. In effect, it represents an engineering approximation, which is why using a model calibration process to arrive at an approximate average value is an acceptable, reasonable, and common approach. Thus, Drs. Spiliotopoulos and Hennet's concern about precisely and accurately defining a value for K_d is misplaced because the theoretical underpinnings for this parameter are not rigorous. That is, conceptual uncertainty in its application must always be recognized, and this conceptual uncertainty carries forward to the use of a conceptually simple retardation factor in the transport equation. This theoretical uncertainty, however, does not preclude the use of these two parameters (K_d and R_f) for characterizing the average transport behavior of a contaminant such as PCE in flowing groundwater.

Zheng & Bennett (2002) describe some limitations in modeling sorption processes. They note that there are significant computational difficulties inherent in coupling advective-dispersive transport with

chemical reactions (p. 79). They further note (p. 79-80) that "… field problems always involve uncertainty as to the nature of the controlling reactions, and as to the quantities and properties of the reacting substances. As a result, the biogeochemical processes represented in field-scale transport models at the present time are largely limited to reactions of the simplest kind, based on highly idealized representations of the effects of more complex reactions."

Kret et al. (2015) studied a Quaternary sandy aquifer to estimate sorption coefficients for PCE fate and transport modeling. They estimated K_d from both batch and column experiments and concluded that reasonable values for R_f for PCE are typically between 1.1 and 3.6.

Rogers (1992) developed a groundwater transport model for the Lawrence Livermore National Laboratory (LLNL) site in California, which includes "several hundred feet of complexly interbedded, unconsolidated alluvial sediments" with an upper boundary represented by an unconfined water table condition. Their calibration and history matching resulted in reasonable matches for R_f values between 1.0 and 3.0, with their conclusion that "a spatially averaged retardation factor of approximately 3 is recommended…"

<u>Model Documentation:</u>

To facilitate assessment of the scientific credibility and scientific defensibility of a groundwater model, the model study should be well documented. Reilly and Harbaugh (2004) state: "Because models are embodiments of scientific hypotheses, a clear and complete documentation of the model development is required for individuals to understand the hypotheses, to understand the methods used to represent the actual system with a mathematical counterpart, and to determine if the model is sufficiently accurate for the objectives of the investigation. … The appropriate level of documentation will vary depending on the study objectives and the complexity of the simulations."

Reilly and Harbaugh (2004) list ten topics that should be addressed in reports documenting model studies. These are:

1. Describe the purpose of the study and the role that simulation plays in addressing that purpose.
2. Describe the hydrologic system under investigation.
3. Describe the mathematical methods used and their appropriateness to the problem being solved.
4. Describe the hydrogeologic character of the boundary conditions used in the simulation of the system.
5. If the method of simulation involves discretizing the system (finite-difference and finite-element methods for example), describe and justify the discretized network used.
6. Describe the aquifer system properties that are modeled.
7. Describe all the stresses modeled such as pumpage, evapotranspiration from groundwater, recharge from infiltration, river stage changes, leakage from other aquifers, and source concentrations in transport models.
8. For transient models, describe the initial conditions that are used in the simulations.
9. If a model is calibrated, present the calibration criteria, procedure, and results.
10. Discuss the limitations of the model's representation of the actual system and the impact those limitations have on the results and conclusions presented in the report.

The documentation for the ATSDR model studies at Tarawa Terrace and HPHB study areas are detailed, comprehensive, and clear, and meet or exceed these guidelines, as evidenced by the series of model documentation reports that include 11 separate book chapters for Tarawa Terrace and 4 separate book chapters and 8 supplemental volumes for HPHB. Careful review of this comprehensive documentation indicates that ATSDR used scientifically acceptable tools and followed correct scientific methodology in performing its historical reconstruction, in contrast to the assertions of Dr. Spiliotopoulos and Dr. Hennet.

Review Comments on Dr. Spiliotopoulos' Opinions:

Opinion 1: Dr. Spiliotopoulos states "Due to the absence of sufficient historically observed data and site-specific parameters, the results of these calculations [in the ATSDR models] are highly uncertain and cannot be used for determining dose reconstructions at the level of detail that ATSDR presented in their analyses." I would counter that although early time data are lacking, there are still a lot of data and historical observations available, as documented in the several ATSDR reports on the investigations. Dr. Spiliotopoulos fails to specify how much data would be "sufficient". In any groundwater modeling study, there are never "enough" data and there is always uncertainty in the final model results. This is normal and expected. In this case, there were enough data to calibrate groundwater flow and transport models, and the data deficiencies were not so great as to prevent a historical reconstruction. In fact, a reasonable historical reconstruction was indeed accomplished, so it was possible. The historical reconstruction recognized the existence of uncertainty and assessed its impact on the results.

Dr. Spiliotopoulos refers to Section 4 of his report as his support for this opinion. Following are comments about his discussion in Section 4 of his report.

In the introduction to Section 4 (p. 27, para. 2), Dr. Spiliotopoulos overstates the lack of data for the Camp Lejeune groundwater system. He says that without site-specific data and a lack of observations, a model "can even be considered speculative and unfounded." That might be true if there were no site-specific data and no observations. But that is simply not true for these models. There are certainly site-specific data available on subsurface properties, as well as observations of heads, boundary conditions, and chemical concentrations for some time periods. These are all described in detail in the numerous reports published by ATSDR. There is no basis for applying the characterization of "speculative and unfounded" to the ATSDR models of TT and HPHB. Even for predictive periods, the system behavior simulated in the model still obeys the laws of physics and hydraulic principles, and contaminants will move in directions predictable by the hydraulic gradient, as calculated with the flow model.

In para. 3 (p. 27), he states that "'predictions' refer to model output, regardless of whether its results are used for hindcasting or forecasting …" I agree with this statement. However, in the next paragraph he discusses "When historical data are not available…" But whether the model predictions are used for forecasting or hindcasting, if it's truly a prediction, then there will be no measurements available (except later for a forecasting prediction). But at the time of model development, observation data for heads and concentrations will only be available during the calibration period. Implying that the lack of data during a predictive period is a problem is misleading. (If data were available during a historical period of interest,

hindcasting would not be needed—it would just be used as part of the observed data set for the calibration period.)

In para. 2 (p. 29), Dr. Spiliotopoulos states that Dr. Clement (in Clement's 2011 publication) "indicated that ATSDR's analysis implied almost exact knowledge of past conditions." I disagree. I find that ATSDR is clear that uncertainty exists about the conditions during the historical reconstruction period, as well as during the calibration period, and the results include assessments of uncertainty. If Dr. Clement inferred that ATSDR believed they had an exact knowledge of past conditions, then that is Dr. Clement's mistake. In the same paragraph on p. 29, Dr. Spiliotopoulos quotes Dr. Clement's comments about the uncertainty analysis. Although the quote starts with Dr. Clement saying that "the results appear to be reasonable …", he ends the quote with an apparent criticism by saying: "The figure also shows that closer to the initial starting point the confidence band is almost 100%, implying that our knowledge of initial conditions, initial source loadings, and initial stresses is almost exact." Although it may be counterintuitive, as I discuss in my Introduction, I actually do have high confidence in the assumption that there were no (or negligible) contaminants in the groundwater from ABC Cleaners prior to Jan. 1953, and probably very little for at least several months after that. Thus, at some point the confidence band should get narrower going backwards in time towards the starting date of the simulation.

In his Summary of Opinion 1 (p. 30), Dr. Spiliotopoulos says "these models were largely not constructed using site-specific data …" I strongly disagree. The geometry and boundary conditions of the model and its hydrogeologic framework are derived from hydrogeologic and geophysical studies of the subsurface aquifer system at the Camp Lejeune and adjacent areas, as documented in USGS reports and in several of the ATSDR reports. This type of information provides a critical and necessary foundation for the models. The potentiometric and water table maps also provide important information for the construction and calibration of the models. Dr. Spiliotopoulos also states in this summary that the models were not "calibrated to observed data for the first 30 years of simulation." Of course, because those concentration data did not exist. That is the reason these models were built—to estimate those concentrations in a state-of-the-art way that is consistent with principles of groundwater flow and transport processes. The models did not generate arbitrary or random numbers. The results are based on the physics of groundwater flow and contaminant transport, and the results appear reasonable and realistic, and the existence of error bands or uncertainty ranges around the estimates is expected and openly acknowledged.

Opinion 2: Dr. Spiliotopoulos says that ATSDR used "parameters and assumptions that are incorrect or not representative of site conditions …" Parameter values for groundwater models are never known precisely and accurately. That is an unfortunate fact of life in groundwater modeling. The parameter estimation process (essentially, the model calibration exercise) is conducted to adjust parameter values within a range of reasonable values to yield a best fit between model simulation results and the limited observation data available. This naturally allows and/or creates compensating errors in the input data for the model. Dr. Spiliotopoulos says this results in conservative estimates of estimated monthly contaminant concentrations. It is not clear what is meant by "conservative" or why that is not a good trait. He also says the results are biased high. His main argument for that opinion seems to be that early (in time) results often lie above the mid-point of the uncertainty bands. The uncertainty bands reflect a zone within which results are expected 95% of the time; if results mostly fall within the uncertainty

bounds, they should be considered acceptable. He cites sections 4.1.1 and 4.1.2 of his report for support of this Opinion.

On p. 31 (Section 4.1, 4th para.) Dr. Spiliotopoulos states "ATSDR's calibrated model sits at the top of the uncertainty range, … This demonstrates that the calibrated model was biased high." But it does not prove ATSDR's model is wrong. The results are within the uncertainty bounds and true values are expected to lie somewhere within the uncertainty bands. Furthermore, best estimates of concentrations do not have to lie at the center of the error band. A model may become insensitive to certain parameters used to create the error bounds at their upper or lower limits, and the response of the model to some parameter variations is not linear.

In para. 7 on p. 31, Dr. Spiliotopoulos quotes the NRC (2009) report where it says "Reporting precise values based on model predictions gives the misleading impression that the exposure of the former residents and workers at Tarawa Terrace during specific periods can be accurately defined." Would he prefer imprecise values? NRC gives no examples of where the ATSDR-reported values are too precise or are prone to misinterpretation in light of the pervasive discussions of model uncertainty provided by ATSDR in its reports. Furthermore, Dr. Spiliotopoulos fails to cite the first sentence of that same paragraph, where the NRC report states "The committee concluded that <u>ATSDR applied scientifically rigorous approaches</u> to address the complex groundwater-contamination scenario at Tarawa Terrace." [emphasis added.]

For Section 4.1.1 (p. 32), Dr. Spiliotopoulos uses the heading "Available data are limited to non-existent", but the first statement after that notes that there were 36 aquifer tests at TT to estimate aquifer properties. This is actually a lot of data, especially considering that aquifer tests are time-consuming and expensive to run. Data for TT are certainly *not* non-existent. I am sure many groundwater models have been developed for areas where there were less than 36 aquifer tests available.

In his summary of Opinion 2 (p. 33), Dr. Spiliotopoulos references his Fig. 5, which includes a reproduction of ATSDR's Fig. F16 about TT results, and goes on to say that ATSDR's work resulted in "biased high estimates." I reproduce that part of Dr. Spiliotopoulos' Fig. 5 (Fig. F16) here because it actually illustrates the opposite. It shows 5 measured PCE concentrations in samples from well TT-26 collected within weeks of each other in early 1985. Over this relatively short time span, the concentrations varied greatly (bracketed between a high of 1,580 ug/L on 01/16/1985 to a low of 3.8 ug/L on 02/12/1985)—a rate of change that cannot be replicated in a model using monthly time steps. Most importantly, the plot shows that the model results fell almost exactly at the midpoint of the range of observed values (about 800 ug/L)—countering the claim of the model being biased high.

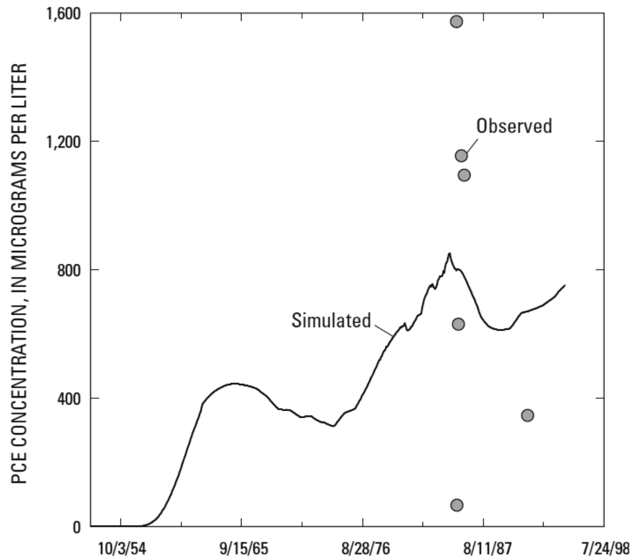

Figure F16. Simulated and observed tetrachloroethylene (PCE) concentrations at water-supply well TT-26, Tarawa Terrace, U.S. Marine Corps Base Camp Lejeune, North Carolina, January 1952–December 1994 (see Figure F6 for location).

Section 4.1.2, p.34, 1st para.: Dr. Spiliotopoulos quotes TT Chapter C (p. C38) saying that "… simulation results are unqualified for the years 1951-1977, …" This is a statement of recognition by ATSDR that there is a paucity of water-level measurements during that early time period. This is also part of ATSDR's consistent messaging that uncertainty exists, and is greater for some time periods than for other time periods. However, it does not disqualify or "unqualify" the model itself, as even during that same time period, other calibration controls and constraints exist in terms of boundary conditions and stresses. Specifically, the adjacent surface water systems represent hydrologic boundaries with known average elevations that change very little over periods of decades. Average monthly recharge can also be estimated based on precipitation and other climatic data that are available. Given such constraints, there is a limited range over which the simulated heads can vary, and that range is not unqualified or unconstrained.

In Section 4.1.2, p.34-36, Dr. Spiliotopoulos cites ATSDR (TT, Chapter F) as noting that 53% of comparisons of simulated to observed concentrations violated ATSDR's calibration target. But many of these samples were collected on the same day or within a short time of other samples (Figure 6 (Table F13), p. 35), so giving equal weight to each comparison is not statistically reasonable. These temporally closely spaced samples are not truly independent samples. Alternatively, I would say a fair comparison should be made on the basis of the quality of the agreement between simulated and observed concentrations at the 11 separate sampling (well) locations. This gives equal weight to every sampling location. Of these, 8 can be deemed "accurate" (including two that have some low and some high samples, so accurate on average), one is high but within the target range, one is slightly high, and one is consistently high (TT-23). On this basis, 73% of the sampling wells show reasonably and acceptably accurate simulation results. Also see my related discussion of calibration targets below (for Section 3.3).

On p. 36, para. 4, in his summary of Opinion 2, Dr. Spiliotopoulos states that the "model calibration did not rely on observed data prior to 1984." Yes, no contaminant concentration data were available then, and that is why ATSDR needed a deterministic groundwater simulation model to estimate how the contaminants were distributed in the aquifer during that time period.

Opinion 3: This Opinion notes that the calibrated model for TT was built using different parameter values and assumptions than the HPHB model. Dr. Spiliotopoulos cites sections 4.1.2.2, 4.1.2.3, and 4.2.3.2 of his report for support. In general, I note that these two study areas do not overlap. Although they are adjacent, and one would expect similar characteristics, having differences is not surprising and certainly the two independent calibrations can yield different values for the various parameters in the models. The models were also developed and calibrated at different times (TT being the earlier model) and improved calibration (parameter estimation) software was applied in developing the latter (HPHB) model.

Dr. Spiliotopoulos (Section 4.1.2.2.1, p. 37) indicates that an error was made in calculating the bulk density (ρ_b) for the TT system. Using an average value for total porosity of about 35%, he calculated that ρ_b should be lower, stating that "In the Hadnot Point model, this error was not repeated." That value was 1.65 g/cm^3. He states that "This has a significant impact on the calculation of the retardation factor, resulting in faster (sooner) arrival of PCE at the water-supply wells, …" However, as Dr. Spiliotopoulos himself admits, this significant impact on R_f does not actually occur because the calibration process compensates for an overestimate of ρ_b by estimating a value for K_d that appears to be too low. Recall that neither of these two parameters are used directly in the transport model. Rather, the retardation factor is used to calculate the migration velocity of the contaminant, and this retardation factor depends on the product of ρ_b and K_d. The calibration process yields a very reasonable value for R_f for PCE—a value (about 2.9) that is very consistent with values in other field studies reported in the literature (e.g., Rogers, 1992; Kret et al., 2015). In Section 4.1.2.3, Dr. Spiliotopoulos has a whole paragraph describing the erroneous consequences "if ATSDR had used a retardation factor of 6.44." But ATSDR did *not* use a R_f = 6.44, so this argument is irrelevant. In summary, the two specific possible errors cited by Dr. Spiliotopoulos for ρ_b and K_d largely offset each other, and have a minimal or negligible impact on the final results, as documented by ATSDR (CLJA_WATERMODELING_01-0000075468; ATSDR_WATERMODELING_01-0000887324).

Dr. Spiliotopoulos (Section 4.1.2.4, p. 39 and elsewhere) and Dr. Hennet (Opinion 11) raise concerns that site-specific data were not used to estimate total organic carbon (TOC) or to calculate K_d. TOC is used to estimate f_{oc}, which in turn is used together with an estimate of K_{oc} to estimate K_d, which in turn is but one factor in the equation used to estimate R_f. That is a long string of dependencies. Appendix A of Dr. Spiliotopoulos' report shows that reported values of TOC vary over a range of about four orders of magnitude. That is a huge variation and uncertainty, which is not accounted for. You cannot simply assume that the mean of that distribution of TOC values is the true and correct one to use to estimate K_d. Overall, there would be much less uncertainty, greater value, and more clarity in just estimating an average value for R_f as part of the calibration process, which is the methodology ATSDR employed. I believe that this is not optional and that R_f must be estimated during and in accordance with the calibration process. In light of this, it simply would not have mattered if K_d had been preliminarily

estimated by ATSDR using highly variable site-specific measurements of f_{oc}/TOC. In the end, the value of R_f = 2.9 calibrated by the ASTDR modeling work is very close to other values reported in the literature for aquifers having similar geologic materials.

Dr. Hennet also criticizes ATSDR for failing to consider available site-specific data for f_{oc} (fraction of organic carbon) to estimate values of K_d (his Opinion 11). Rogers (1992, p. 51) in discussing the K_d parameter says "Numerous researchers have used theoretical methods correlating the organic carbon content (OCC) of the subsurface material and the K_d (Karickhoff, 1984). Others have used the partitioning between octanol and water to predict the K_d (Kenega, 1980). **These methods are not considered appropriate where the OCC is less than approximately 0.1%**." OCC is equivalent to TOC, and 0.1% is equivalent to a fraction or 0.001. Hennet's Expert report lists (Exhibit 3-2, and p. D-11 to D-12) 21 Camp Lejeune samples where f_{oc} is given. The median value is 0.0013, barely above the indicated limit, and 9 samples (43% of the samples) have values <0.001, indicating that the use of f_{oc} to estimate K_d is not appropriate. If ATSDR had used this approach, it would have introduced additional errors and sources of uncertainty.

In his summary of his Opinion 2&3 (p. 38-39), Dr. Spiliotopoulos states (in reference to ρ_b and K_d) that "parameter values in the Tarawa Terrace model were different than those used in the Hadnot Point model, even though both models simulated similar hydrogeologic conditions." This is not a problem, and it would be more surprising if they had applied identical values. The areas have similar conditions, not exactly the same conditions. Hydraulic conductivity measurements show notable differences between the two areas, reflecting local differences in aquifer material properties. These differences also cause differences in the factors contributing to the R_f. There is nothing wrong or unexpected about this. R_f was estimated in the calibration process, and the HPHB calibration used a different (and supposedly better) automated parameter estimation software package, which was not used in the TT calibration. So of course some differences will result. If they had applied the same parameter estimation software to both sites, it still would most likely result in different values for the average R_f in the two different areas. But the differences are small and inconsequential.

In a summary of his Opinion 3 (p. 39), Dr. Spiliotopoulos states that "these incorrect assumptions resulted in faster plume migration in the aquifer and estimated monthly concentrations that were conservative and biased-high." However, this would only be the case if the errors in the two parameters were considered separately and alone. But the model does not respond to these values separately. It responds to their net effect on the retardation factor, which was calibrated to a very reasonable value consistent with other peer-reviewed studies. The errors were compensatory and that compensation was built into the critical R_f value by the calibration process, as would be expected from a calibration process for a groundwater model.

<u>Opinion 4</u>: Dr. Spiliotopoulos says that use of "parameter values based on site-specific data … in Tarawa Terrace would result in substantially lower estimated monthly concentrations. Furthermore, the model uncertainty range would also be lower." Dr. Spiliotopoulos cites his Section 4.1.2.5 as support.

On p. 39, Dr. Spiliotopoulos argues that site-specific data for calculating K_d would result in a higher K_d value. Again, the model calibration process adjusted values of K_d, one component of the retardation

factor, so that the value of R_f was as reasonable and accurate as possible for maintaining consistency with the available observed concentrations. Furthermore, in calculating K_d, Dr. Spiliotopoulos used a porosity value of 20%, which was the effective porosity used in the transport model. However, in calculating ρ_b, the other component of R_f, Dr. Spiliotopoulos used a porosity value of about 35%--a value representing the total porosity measured in two soil samples (p. 37). Using two different values for porosity in the same equation is inherently wrong, creating an inconsistency of 75%, and is done with no explanation.

In section 4.1.2.5, Dr. Spiliotopoulos develops a "revised" model using a late start date and a different K_d value. He presents his results in comparison to the ATSDR model results in his Figs. 7 and 8. He accentuates the early time differences by plotting results arithmetically rather than logarithmically. But that's a minor point. The proper start date is outside the scope of my opinions. But adjusting the K_d without also adjusting p_b is one-sided. In any case, Dr. Spiliotopoulos' value for R_f in the revised model is 3.48. The value of 2.93 used by ATSDR is only 16% lower than this new value used in Dr. Spiliotopoulos' revised model. This difference is relatively small. Furthermore, as seen in those two figures, the difference between the ATSDR results and Dr. Spiliotopoulos' revised model results are very small after approximately 1970. More importantly, both models are consistent in showing that PCE concentrations are above the MCL for most of the study period—and since Jan. 1, 1960 in both models, at both Well TT-26 and in influent to the TT WTP.

Also noteworthy in Dr. Spiliotopoulos' Fig. 7 is that for both models, there is a peak concentration shortly before 12/84. When K_d is higher and R_f is consequently higher, then one would expect that a peak moving through the groundwater system would be somewhat delayed, yet there is no indication in the results for Dr. Spiliotopoulos' revised model that this peak concentration was delayed at all. Instead, it appears to have arrived at TT-26 at the same time as in the ATSDR model. This demonstrates a lack of sensitivity to the value of K_d in this particular system. It simply did not make a significant difference.

Dr. Spiliotopoulos' only support for his opinion that the uncertainty range would be lower is a concluding statement in his Summary on p. 41, which states, "The uncertainty range for such historical reconstruction would also be lower, as it would be based on slower plume migration and lower concentrations for many years after the start of contaminant releases from the source." However, this is an inference that itself is not supported by analytics. Dr. Spiliotopoulos has not demonstrated that the uncertainty range would be lower. Dr. Spiliotopoulos' results also do not demonstrate significantly slower plume migration (peaks are coincident) or significantly lower concentrations (after 1970 they are almost identical—differing at TT-26 by an average of about 30 or 40 ug/L out of an average concentration of roughly about 500 ug/L—less than 10%).

Opinion 5: This opinion states that the ATSDR groundwater model for TT "resulted in biased-high estimates of monthly contaminant concentrations at one of the water-supply wells." The well in question is TT-23. Dr. Spiliotopoulos cites Section 4.1.2.6 of his report in support of this opinion.

Section 4.1.2.6 (p. 42) offers no clear evidence that the discrepancy at this one well (out of many) has a substantial impact on the overall results. Based on ATSDR Table E2, of the nine unique sampling dates for this well, six had an observed level of PCE or TCE above the MCL. Furthermore, with respect to the overall effect on concentrations estimated at the WTPs, it is important to note that TT-23 was

operational for only about 9 months or less, starting in 1984, and had the shortest operational (pumping) period of any of the 16 pumping wells operating in the TT area (see Table H3 in Chapter H of the TT series of reports). When it was pumping, the contribution from this well provided only a small fraction of the total groundwater inflow to the WTP with concentrations far less than well TT-26 (with its modeled concentrations likely being underestimated). Thus, if indeed the estimates for this well were too high (by less than two times), the effect on calculated concentrations in the WTP would be minimal both in magnitude and in duration.

<u>Opinion 6:</u> Dr. Spiliotopoulos says that the ATSDR model did not reflect "observed data that indicated absence of contamination in the aquifer." Does he doubt that there was contamination in the aquifer? The presence of contamination in the aquifer is well documented; the absence of contamination in some locations means little overall—only that the contamination was not everywhere. That is normal. The statement and implication that there is no contamination in the aquifer is simply incorrect. The ATSDR reports clearly document observations where the contaminants were not detected (e.g., Table F13), and their analyses reflect that. Support for this opinion is stated to lie in Section 4.1.2.7.

In Section 4.1.2.7, Dr. Spiliotopoulos makes a major point about plotting non-detects, and he criticizes ATSDR for not plotting nondetects. He cites the reason as being that "non-detections listed as zeros are not visible in a logarithmic-scale scatterplot. This is because a logarithmic scale can only show numbers greater than zero." However, nondetects do not mean that the value is zero—only that it is less than the detection limit. In aiming to support his point, Dr. Spiliotopoulos relies on an analysis that is arbitrary, incorrect, and biased. He selects a value of 0.1 ug/L to represent all nondetects. For these samples, the detection limits were between 2 and 10 for most analyses. Helsel and Lee (2006) say: "The most common procedure within environmental chemistry to deal with nondetects continues to be substitution of some fraction of the detection limit. This method is better labeled as "fabrication", as it reports and uses a single value for concentration data where a single value is unknown. Within the field of water chemistry, one-half is the most commonly used fraction, so that 0.5 is used as if it had been measured whenever a <1 (detection limit of 1) occurs." If representing nondetects in a plot is to be done, a reasonable value and common way to represent a nondetect would be halfway between the detection limit and zero. For the Camp Lejeune data with detection limits of 2.0 and 10.0, the plotted position should be either 1.0 or 5.0 respectively (the latter being 50 times greater than the arbitrary value Dr. Spiliotopoulos used—so plotting 0.1 instead of 5.0 is a significantly misleading/biased-low way to present the data). This will make a big difference on his plot (such as his Fig. 18). Note: On this topic, Helsel and Lee (2006) also state: "All such plots [scatterplots using halfway points] are misleading, because unique censored values are unknown. Instead, left-censored data can be plotted as intervals between zero and the detection limit for each observation. In this way, no false statements about where an individual value is located, or that all such observations are at the same value, are made." There may also be other alternatives for plotting nondetects (newer and better, but more complicated). Regardless, Dr. Spiliotopoulos' selection of 0.1 to represent all nondetects is arbitrary, misleading, and wrong. ATSDR's approach of not plotting nondetects avoids the possible perception of "fabrication" and is more defensible than Dr. Spiliotopoulos' approach of assuming all nondetects can be fairly represented by an arbitrary value of 0.1, as shown in his Fig. 9 (p. 43). The discussions of Helsel and Lee (2006) justify the ATSDR's approach for not including nondetects on the data plots because of the risk of appearing to

fabricate data or presenting misleading plots. ATSDR does show nondetects in all tables of measured concentrations.

In para. 1 (p. 45), Dr. Spiliotopoulos notes that the model results indicate a low value of 5.8 ug/L in well TT-54, but the observed value was a nondetect. He states that the calibration "is not supported by the non-detection in the sample collected in February 1985." I would argue that it is indeed supported by that data. The detection limit for that analysis was 10 ug/L (TT Table F2). The halfway point between zero and the detection limit is 5.0, a value that is very close to ATSDR's simulated value, and that close agreement is certainly supportive of the quality of the calibration.

Dr. Spiliotopoulos notes (p. 45) that "Well TT-54 had a reported non-detection in July 1991. However, the ATSDR model indicated an increasing concentration trend at well TT-54, suggesting that the PCE plume continued arriving at that well until that time. This is unlikely to be accurate." However, if one examines the predevelopment and transient potentiometric surfaces (TT Chapters C and F), it is clear that TT-54 is downgradient from the ABC Cleaners, and that a plume evolving from that source while several water-supply wells are operational will likely contribute some contaminants to well TT-54.

Dr. Spiliotopoulos' Summary of Opinion 6 (p. 45) picks two of the wells to generalize that "ATSDR's model overestimated the plume migration extent and rate of migration, which were both conservative and biased-high." This is an overgeneralization that ignores other wells and locations where estimates were very close or were underestimated. The nature of model calibration is that there will be compensating errors and that some simulated values will be too high and others too low. Certainly, the results for the flow model (e.g., Fig. C9) do not support a generalization that the flow model is inaccurate or biased-high.

<u>Opinion 7:</u> Dr. Spiliotopoulos states that "the presentation of results of the uncertainty analysis conducted by ATSDR for the Tarawa Terrace model was misleading by showing a narrow uncertainty range around the calibrated model." Support is given in Section 4.1.3.1.

In 4.1.3.1, Dr. Spiliotopoulos' characterization changes from "misleading" to "visually misleading." The stated reason is that "they used a logarithmic scale, which visually compresses the uncertainty range around their calibrated model [results]." However, the use of a logarithmic scale is a valid and common approach in engineering and scientific studies, and is not characterized as being misleading by scientists and engineers. He observes that the plot ranges over six orders of magnitude on the axis for PCE concentration, but the width of the uncertainty bands do not. When values span such a large range, it is normal and standard to use a log plot. Using just an arithmetic scale would effectively hide all the changes in the lower part of the scale.

Dr. Spiliotopoulos states (p. 46, para. 4) that "the difference between the high and low values in Figure 11 [ATSDR's Fig. I29] is not significant enough to justify the use of a logarithmic scale." I disagree because the observed values span more than two orders of magnitude (excluding nondetects) and the simulated values span more than five orders of magnitude. Plotting these using a log scale is reasonable and informative, and is the only way to portray the early time results of the simulation in the same graphic. It is fine to also present these results plotted on an arithmetic scale (Fig. 12), but not sufficient to do so solely. Dr. Spiliotopoulos' concern over the concentration plots is mostly cosmetic.

On p. 48 (para. 1), Dr. Spiliotopoulos criticizes the uncertainty analysis, saying "… the concentrations calculated by the model should be generally in the middle of the uncertainty range … However, the calibrated model-simulated concentrations are almost identical to the upper bound of the uncertainty range in the early years of operation (1957-1963)." However, if one examines his Fig. 12 (p. 48 of his report), it clearly shows that the results are indeed generally in the middle of the uncertainty range. In the few early years it is above the middle, but consistently below the upper bound, as desired. Such a result is within a probabilistic expectation. In those early years the concentrations are the smallest. For example, in 1960 the difference between the upper bound and the middle of the range is only about 10 ug/L, which is a small value on the full scale of PCE values considered. Being "generally near the middle" is not an objective or quantitative rule.

Opinion 8: Dr. Spiliotopoulos states that "ATSDR's uncertainty analysis was not bound by historical concentration data, and as a result, focused only on model precision and not accuracy in predicting COC concentrations. ATSDR's uncertainty analysis was presented as though it evaluated the model's accuracy. It did not." Support is stated as being in Section 4.1.3.2.

The criticism is based on the lack of historical data on concentrations prior to 1982 (Section 4.1.3.2, p. 49), and would mean that "the uncertainty analysis would result in precise but not necessarily accurate solutions …" However, once again, the lack of concentration data prior to 1982 is the reason that the model was developed. Data are available afterwards, and initial conditions for the contaminant distribution can be stated with reasonable reliability that the concentrations in the TT area were zero prior to the start of operations at ABC Cleaners. That is an important known concentration condition for the early 1950s. What the model does is estimate how the concentration changed spatially between the time of the start of ABC operations and the time when observations of PCE became available, and it does so in a manner that is consistent with the principles of groundwater flow field and solute transport, with the further recognition that the groundwater flow field has been simulated with acceptable accuracy.

The ATSDR assessed uncertainty using a sophisticated but standard and acceptable statistical approach—using a Monte Carlo simulation method. They carefully documented their approach, which generated 840 realizations. In a Monte Carlo simulation approach, no single realization is expected to be "accurate." Rather, the ensemble of realizations is intended (and expected) to bracket a range of feasible but realistic outcomes. The range of results (generally considering 95% of the outcomes) is a measure of the model's predictive accuracy. The Monte Carlo uncertainty analysis would not be expected to yield a different calibrated model.

In the last paragraph on p. 49, Dr. Spiliotopoulos states that "one of the most critical parameters for determining how fast contaminants will migrate in the aquifer is the retardation factor." I would argue that both the speed and direction of migration is more critically determined by the head distribution (hydraulic gradients, as determined by the groundwater flow model) and the effective porosity. The retardation factor will have no effect on the direction of transport of a contaminant for a given flow field. Furthermore, the results presented by Dr. Spiliotopoulos in his Fig. 7 show that the model results, at least at Well TT-26, are relatively insensitive to a range of uncertainty in the assumed value of K_d and R_f.

On p. 50 (para. 3), the Monte Carlo approach used by ATSDR is criticized by Dr. Spiliotopoulos "… because ATSDR implemented a 'probability distribution function' … to describe how values closer to the mean

value of the range are more probable than those away from the mean." I do not see a problem here as this is an option within standard practice for random sampling of parameter values for a MC analysis when information or theory indicates that a parameter has a statistically normal or log-normal distribution. Zheng & Bennett (2002, p. 353) say "The Monte Carlo method is by far the most commonly used method for analysis of uncertainty associated with complex numerical methods." They further state (p. 356) "The heart of the Monte Carlo method is the generation of multiple realizations (or samples) of input parameters that are considered to be random variables. Each random variable is assumed to follow a certain probabilistic model characterized by its probability density function (PDF). The probability distributions commonly used in hydrogeologic studies include *normal*, *lognormal*, *exponential*, *uniform*, *triangular*, *Poisson*, and *beta* distributions." It is worth noting that when this book was published, co-author Bennett was an employee of SSP&A and first author Zheng was a former employee and affiliate of SSP&A.

The plots shown in Fig. 13 are discussed in para. 8 (p. 50, Section 4.1.3.2). Dr. Spiliotopoulos notes that the results of the calibrated model "sits at the upper bound of the retardation-factor uncertainty range." However, that is not true for the majority of the simulation period. It is close to the middle of the range during the period of 1962 through the end (around Dec. 1987). And prior to 1962, it still lies within the uncertainty bounds, which is acceptable and not indicative of bias. As stated earlier, error bounds need not be evenly distributed around the mean because a model can be sensitive to a parameter at either high or low values, but not both.

In the 3rd paragraph on p. 51, Dr. Spiliotopoulos presents the values for the retardation factor with four significant figures. Whether R_f is estimated by adjustments during model calibration or estimated from highly variable and uncertain site-specific data, presenting it with 4 significant figures is an unjustified and meaningless precision.

Opinion 9: This continues the previous discussion of the uncertainty analysis and cites the same section (4.1.3.2) as support. Dr. Spiliotopoulos says that the uncertainty analysis for TT "… did not encompass uncertainty bounds representative of site-specific conditions, resulting in biased-high uncertainty range."

It is not clear exactly what is meant by a "biased-high uncertainty range." If it means that the uncertainty range is incorrectly too high, that implies that the model is even more accurate than indicated.

On p. 52 and in Fig. 14, Dr. Spiliotopoulos discusses the results if R_f were 4.3 instead of 2.9. But this value of 4.3 is higher than those presented in published peer-reviewed articles of PCE transport in similar types of aquifer materials (Rogers, 1992, and Kret et al., 2015). Even with Dr. Spiliotopoulos' high value of R_f, Fig. 14 shows that after about 1970, the differences at Well TT-26 are small—less than 100 ug/L difference during the final 20 years of the simulation, with Dr. Spiliotopoulos' revised model showing lower concentrations because it includes a larger sorption rate. Again, it is relevant to note that the observed data shown in this figure range from about 3 ug/L to almost 1600 ug/L for samples collected over a relatively short time period in early 1985. The ATSDR model results fall very close to the midpoint at that time—at about 800 ug/L—not indicative of any bias. However, Dr. Spiliotopoulos' revised model with the higher R_f value calculated a PCE concentration of about 700 ug/L at the time when the data are available—lower than the mid-point, which does not provide evidence that the higher value of R_f is more

accurate (actually, it's an indication that it is less accurate). Either way, the computed PCE concentration values are higher than the MCL for all times after 1960, which is a critical point.

The three highest observed values of PCE in well TT-23 were underestimated by the ATSDR model, which counters the claim that the ATSDR model is biased high.

On p. 55, Dr. Spiliotopoulos says that "ATSDR's selection of the retardation factor parameters forced the calibrated model to simulate fastest arrival of PCE at well TT-26 …" This use of the word "forced" appears to unfairly attribute an unscientific and biased motive to the way the model calibration was conducted. First of all, this was not the fastest possible arrival. If they had used a value of $R_f = 2$, the arrival would have been faster than the value they calibrated to. I think a fairer way to characterize the calibration relative to R_f is that they varied the values of R_f and of other parameters and selected parameter values that yielded the best overall fit to the available data. This happened to be a value of 2.9 for R_f, which was very consistent with other values reported in the literature for PCE transport in similar types of geologic material.

<u>Opinion 12:</u> This opinion focuses on the model post-audit performed by Jones and Davis. The opinion says that the post-audit showed that "ATSDR's dose reconstruction groundwater model for drinking water in Tarawa Terrace used parameters and assumptions that resulted in conservative and biased-high estimates of monthly contaminant concentrations." Support is said to be given in Section 4.1.5.

It is my understanding that Jones and Davis, as well as Maslia, will respond to this opinion in their rebuttal reports. A few general comments about the content of section 4.1.5 follow.

In Section 4.1.5.1 (p. 60, para. 2) Dr. Spiliotopoulos states that "Observed concentrations of zero correspond to non-detections." As mentioned previously, this statement is not accurate in the sense that nondetect values do not necessarily have a value of zero, but their value may be anywhere below the detection limit for that particular analysis. Also, in para. 3 and Fig. 18 (p. 60), Dr. Spiliotopoulos repeats the same error in assuming that a nondetect can be substituted by a value of 0.1 ug/L. This is arbitrary and biasing.

Dr. Spiliotopoulos calculates a mean error for partitioned segments of the data set—separately for points where the observed value is higher and separately for points where the simulated value is higher. This is not a common or standard way to compute a mean error. Based on my experience and expertise, the standard methodology is to compute the mean error for all data.

<u>Opinion 13:</u> This opinion also focuses on the model post-audit performed by Jones and Davis, and is closely related to Opinion 12. It suggests what Maslia and Aral should have done with the data of Jones and Davis. Support is again said to be given in Section 4.1.5. It is my understanding that Maslia will respond to this opinion in his rebuttal report, but I have a general comment regarding the absence of data.

On p. 63, Dr. Spiliotopoulos expresses concern that "no data are available to evaluate whether the overall extents of the simulated plume are real." Some data are certainly available. It would be nice if

more data were available. If extensive data were available to map the plume in detail over time, there would be little need for a simulation model. The ATSDR models reliably simulate the groundwater flow field and head distributions so that the transport models can simulate advective and dispersive processes, as modified by chemical reactions and adsorption (as simplified using the retardation factor), to fill in the gaps in the observational database in a way consistent with widely accepted governing principles of groundwater hydraulics and transport phenomena. This is a reasonable and appropriate approach to addressing this issue.

<u>Opinion 14:</u> This opinion restates previous ones, but for Hadnot Point, and says that the ATSDR model "was constructed and calibrated using parameters and assumptions that are uncertain or incorrect." Support is said to be given in Sections 4.2.1, 4.2.2, 4.2.3, and 4.2.4.

In general, groundwater systems occur within subsurface geologic frameworks that are complex, heterogeneous, and hidden from view. There are and always will be uncertainty associated with even the best efforts to define the properties and relevant characteristics of these systems. This does not preclude the development of reliably sound numerical models to simulate groundwater flow and transport processes. But model developers must always be aware of, and assess, the existence of uncertainty and the sensitivity of the model results to this uncertainty. ATSDR has indeed accomplished this. For TT, they have produced a 187-page chapter (Chapter I) solely about this task (in addition to many discussions of it throughout the other chapters). For HPHB, there are two sections in Chapter A of their reports focused on these topics.

Dr. Spiliotopoulos states (p. 68, para. 4) that "Unlike the Tarawa Terrace model, ATSDR did not know the precise location of all contamination sources and the magnitude of contamination each source contributed." This is true—there is uncertainty in the source terms (as with all model parameters). But that can be handled and does not preclude the development of a reasonable flow and contaminant transport model. Assumptions had to be made, but they were not "arbitrary" and were clearly and comprehensively documented. He cites the NRC (2009) report, which said "There were multiple sources of pollutants, including an industrial area, … [etc.]" What is certain is that all of these are likely sources of groundwater contamination. Industrial operations in the 1950s, 60s, and 70s were typically not concerned with protecting groundwater quality.

In footnote 235 (p. 68-69), Dr. Spiliotopoulos says, "ATSDR used simulated contaminant concentrations in the influent to the WTP to calculate concentrations in the water delivered to a family housing or other facility, without considering any contaminant losses during treatment." However, unless the treatment process was designed to treat these contaminants, it would have been "arbitrary" and highly uncertain to simply assume that the treatment reduced contaminant concentrations or removed contaminant mass.

p. 69: Dr. Spiliotopoulos cites "evaporative" losses in a treatment plant. However, evaporation is rarely significant in a water treatment plant and direct evidence would be needed to support this hypothetical claim. Contaminant loss due to volatilization during the treatment and distribution process was discussed at the March 28, 2005 expert panel meeting where panelists—including Dr. Pommerenk of AH Environmental— opined that any loss would be minimal (See March 28, 2005 Expert Panel Meeting Transcript at 55:2-57:14, 56:2-57:14).

In para. 3, Dr. Spiliotopoulos says "Based on [his] professional judgment, there was insufficient data to conduct groundwater flow and contaminant transport model calibration and uncertainty analysis." But in fact, ATSDR did "conduct" it, and clearly documented their calibration and uncertainty analyses. In my professional judgment, they did a good job with the limited data available.

In para. 4 (p. 69), Dr. Spiliotopoulos repeats that "prior to 1982, no water quality data were available …" However, groundwater flow directions can be deduced with typically small uncertainties, and flow rates (velocities) and advective-dispersive transport can be simulated with some additional uncertainty, but these key processes are reasonably well defined. Also, it is highly certain that prior to the start of these industrial and landfill operations, the contaminant concentrations were zero—an important early-time data point.

In para. 7, Dr. Spiliotopoulos quotes NRC (2009) as saying "simpler modeling approaches should be used to assess exposures from the Hadnot Point water system." While this is easy to say and sounds appealing, they don't say how to do that or what simple modeling approach would work. How does one know if a model is too simple? What processes should be eliminated in the simpler model? In fact, the way to produce a simpler model is to first develop and calibrate a maximally realistic "complex", detailed, and comprehensive model that can be then used to assess which processes or factors have little effect on the results and so can be safely eliminated to produce a simpler model. The benefit cited by NRC is faster and more efficient modeling, but that potential benefit is not a major need here, and the use of models that might be too simple is offset by their reduced realism and risk of oversimplification.

On p. 70 (section 4.2.1), Dr. Spiliotopoulos says "available data are limited or non-existent" but in the first bullet point states that "more than 200 aquifer and slug test analyses" exist. This is a lot of data! There are many groundwater models that have been developed and calibrated on the basis of much fewer hydraulic testing at the specific site of interest.

On p. 70, Dr. Spiliotopoulos is also concerned that pumpage data for individual wells were estimated on the basis of "ancillary data." This is common standard practice in groundwater modeling, as pumpage measurements for wells are often not available or are of questionable quality.

In the last para. (p. 70) Dr. Spiliotopoulos notes that the HP WTP was built in 1942 and during its first 40 years of operation, there were no water quality data for the contaminants of concern. This is unfortunate, but not unexpected; it is rather common for groundwater contamination problems that a chemical that turns out to be problematic at a later date is not monitored prior to that awareness. This is why ATSDR had to use modeling to help reconstruct the historical record as well as possible, using documented quantitative methods. Of course, there will be uncertainty in the results, but they seem reasonable given the information that is available.

p. 71, Fig. 25 (ATSDR Fig. A18): Dr. Spiliotopoulos presents four plots of simulated and observed TCE concentrations at four wells in the HPHB study area. All four plots show that the simulated values were either close to the middle point between observations (HP-602 and HP-608) or below the observed values (HP-634 and HP-601/660). There is no indication here that the model overestimated concentrations (or was biased-high).

In summarizing Opinion 14 (p. 71), Dr. Spiliotopoulos says "Selection of model parameters was based, primarily, on professional judgment." This is always the case. Data are always limited, and professional

judgment is required to assess how to deal with that paucity of data and how much weight to give the limited number of measurements. A groundwater modeler always wishes they had more data, but the reality is that there are never so much data available so as to avoid using professional judgment.

In Section 4.2.2 (p. 72) the claim is made that ATSDR "made arbitrary assumptions to reconstruct pumping history …" In my opinion, the assumptions were not arbitrary, but rather were well-informed, well-reasoned, and carefully documented. Assumptions had to be made about the pumping history, and they were made, but they were not arbitrary. For example, Dr. Spiliotopoulos notes that "Yearly volumes are available for some years prior to 1980. A trendline was used to estimate raw-water flows for years prior to 1980 when no data exist." This appears to be a sound statistical approach, and the use of a trend line is certainly not arbitrary.

In Section 4.2.2 (p. 72-73) Dr. Spiliotopoulos offers a further criticism that "it was assumed that a well would be operated in the historical period based on a pattern similar to the more recent 'training period,' with further adjustments to account for information on the varying capacity of wells, where available." Dr. Spiliotopoulos' statement actually contradicts his assertion that estimates were arbitrary. Here he describes a reasoned and reasonable approach to estimating a pattern of past water use (well pumpage)—an approach that is not "arbitrary."

In several additional paragraphs on p. 73 (as well as elsewhere), he repeats the claim that pumping rates were based on arbitrary assumptions. ATSDR uses sound statistical methods (such as regression and correlation) to estimate pumpage. This is neither arbitrary nor unreasonable. Similar wells managed by the same operating authority are likely to have been operated in a similar manner. If not, that would be arbitrary. It is unlikely that Dept. of Navy engineers operating the well fields did so in an arbitrary manner. In the early years they just weren't required to maintain as detailed records as would be expected today. Again, ATSDR made reasonable assumptions with the data that they had available.

Near the top of p. 77, Dr. Spiliotopoulos states that model calibration was "improperly influenced" by "erroneous concentrations reported for well HP-634 … while non-detections were ignored." It has not been established nor agreed that erroneous concentrations (actually, one single value) were reported for well HP-634. This is discussed in more detail below in reference to Section 4.2.3.3. Non-detections were not ignored. They are clearly listed and labeled in many tables presented in the ATSDR reports (such as Table A4 in Chapter A of the HPHB report series, and in many other places too).

In Section 4.2.3.1 (p. 77) Dr. Spiliotopoulos claims that "The groundwater flow model has significant limitations in the absence of data for calibration." Although the model has limitations, there is no evidence that the limitations are significant for the purposes that the model was developed. Furthermore, there is not an "absence of data for calibration." In the very next paragraph, Dr. Spiliotopoulos notes that more than 700 water-level measurements were used in calibrating the predevelopment model, which is also the initial conditions for the transient groundwater flow model. Also, there are a lot of data available on the boundary conditions and hydrogeologic framework for the model.

In the 6th paragraph (p. 77), Dr. Spiliotopoulos indicates that the simulation of contaminant transport in the aquifer is inherently uncertain. This is true for all groundwater models. But the uncertainty does not mean that the model is not useful.

In Section 4.2.3.2, p. 78, Dr. Spiliotopoulos notes that ATSDR recognizes that explicit data defining source locations and mass loadings are not available, but then he criticizes ATSDR by saying "these quantities were arbitrarily assigned to the model in order to fit the limited water-quality data available starting in 1982." However, by criticizing ATSDR's methodology, Dr. Spiliotopoulos in effect is criticizing the essence of the model calibration, history matching, and parameter estimation process practiced in groundwater modeling, in which parameter values are adjusted (either manually or automatically) in order to improve the fit (e.g., see Hill and Tiedeman, 2007). Furthermore, the source locations and mass loadings were not "arbitrarily assigned." The general locations of the sources are well-documented, and sources were placed in the vicinity of these documented locations. Consistent with principles of model calibration, the exact placement and strength of these sources were varied within limits until the observed concentrations were reasonably matched by the model. The variation in the exact location, timing, and strength of sources is rarely known, and adjustment of source properties is a commonly-accepted part of calibrating a flow and transport model.

p. 79: Dr. Spiliotopoulos discusses the lack of data to define the source loading terms for the model in the HPHB area. However, there is no doubt that these chemical contaminants (including TCE and PCE) were present in the groundwater at toxic concentrations (above the MCLs) in the HPHB area, and that they were pumped out of the aquifer by several operating water-supply wells.

p. 79: In the summary for Opinion 14, Dr. Spiliotopoulos criticizes the ATSDR for having "assumed constant mass loading of the same magnitude at all sources for more than 40 years", which he characterizes as "highly uncertain, if not impossible." Viewed from a different perspective, what ATSDR did was apply an average rate over the critical time period because there was no basis for differentiating how the loading might have varied over time. In my opinion, this was a reasonable approach. Furthermore, the constant source resulted in a reasonable model calibration, and so there was no reason to incorporate a variable source in the absence of data on transient source characteristics.

Opinion 15: In this opinion, Dr. Spiliotopoulos repeats the claim that ATSDR included an erroneous value in its analysis and model calibration (presumably for the 1,300 ug/L value measured in a sample from HP-634). Section 4.2.3.3 is cited for support.

In Section 4.2.3.3, Dr. Spiliotopoulos argues that concentration data for well HP-634 was incorrectly interpreted and that the reported value of 1,300 ug/L on Jan. 16, 1985 "should be considered erroneous" (although he considers other samples from that well that showed non-detects to be valid). I believe that his basis for this conclusion is speculative and unsupported by facts, as discussed below.

On p. 80, Dr. Spiliotopoulos says "it is unlikely that this well [HP-634] was ever contaminated with elevated TCE concentrations," and he and Dr. Remy Hennet argue that the analysis showing a concentration of 1,300 ug/L should be thrown out. Although Dr. Spiliotopoulos and Dr. Hennet claim the well was shut down permanently, documentation suggests that HP-634 was online in January 1985 (see CLJA_CLW00000004559, CLW4546, and CLW1818). However, even if the well was shut down permanently shortly before the date this sample was collected, I strongly disagree with Dr. Spiliotopoulos' argument that "contamination could not have reached that well when it was non-operational." It is plausible and possible that TCE could have reached the well sometime after the previous sample had been collected. As Dr. Spiliotopoulos surely knows, after a pumping well is shut off,

water levels do not instantly recover and the head distribution does not instantly return to a nonpumping configuration and nonpumping hydraulic gradients. During predevelopment (nonpumping steady-state) conditions, flow near HP-634 is predominantly to the west and southwest (see HPHB reports Fig. A19 for 1951, reproduced below). While this well was operational, a cone of depression (a drawdown of water levels) formed around it, lowering the heads and reversing local hydraulic gradients, and enabling the movement of contaminants from nearby areas containing contaminants west of HP-634 to move eastwards towards HP-634 (as also shown for later times in Fig. A19 below). When a well is shut down, the heads take time to recover (recovery is not instantaneous). During the slow recovery period, water and contaminants will continue to move towards the well while the cone of depression is slowly filled in and recovers. This simple normal response of groundwater systems to the cessation of pumping easily explains the presence of contaminants in a sample collected after the pumping was stopped. Note that concentrations of DCE and VC were also unusually high in this same sample, so the TCE value is not an isolated "outlier" (see table C7 in report Chapter C). This progression is seen in the maps for all three layers for the November 1984 maps shown in Fig. A19 below, where the contaminant is shown to have moved very close to HP-634 from its previous location in the industrial area just to the west. If Dr. Spiliotopoulos argues that it is not possible for contaminants to reach HP-634 once its pump ceases operation, then it is contingent on him to provide some evidence that (a) the recovery is so fast that it is irrelevant (i.e., how long would it take for the hydraulic gradients to reverse again and return to a predevelopment condition?), and (b) that the contaminants were so far from HP-634 when it was shut off that it could not have migrated that distance during the recovery time. Without such calculations or evidence, one can conclude that it is indeed possible for contamination to reach that well shortly after it became non-operational. The primary evidence that it did become contaminated is the measurement of 1,300 ug/L in the January 1985 well sample, and I do not see conclusive evidence that that sample analysis should be discarded.

Dr. Hennet argues that this well was not contaminated by TCE because some vials in the shipment were broken (he does not say the samples for this analysis were in broken vials, so the relevancy of other vials being broken is not apparent). I doubt that the lab would or could perform an analysis or report a value on a sample taken from a broken vial. Dr. Hennet says a CCLJ report shows the value as 10 ug/L. However, the lab that did the analysis reported 1,300 ug/L. Hennet and Spiliotopoulos also say that the value of 1,300 is an outlier, so should be discarded. But there are many high-valued "outliers" in the record, and the record shows other instances where the value can change over similar large magnitudes in a very short time (e.g., TT-26 shown in Fig. F16, where the PCE concentration changed from 1,580 to 3.8 ug/L in successive samples collected just 4 weeks apart, mirroring the change in HP-634 from ND to 1,300 ug/L in a similar 4 week timeframe). The reasoning by Dr. Spiliotopoulos and Dr. Hennet to discard this reported value seems entirely speculative. They offer no actual evidence that the analysis or its reporting was erroneous.

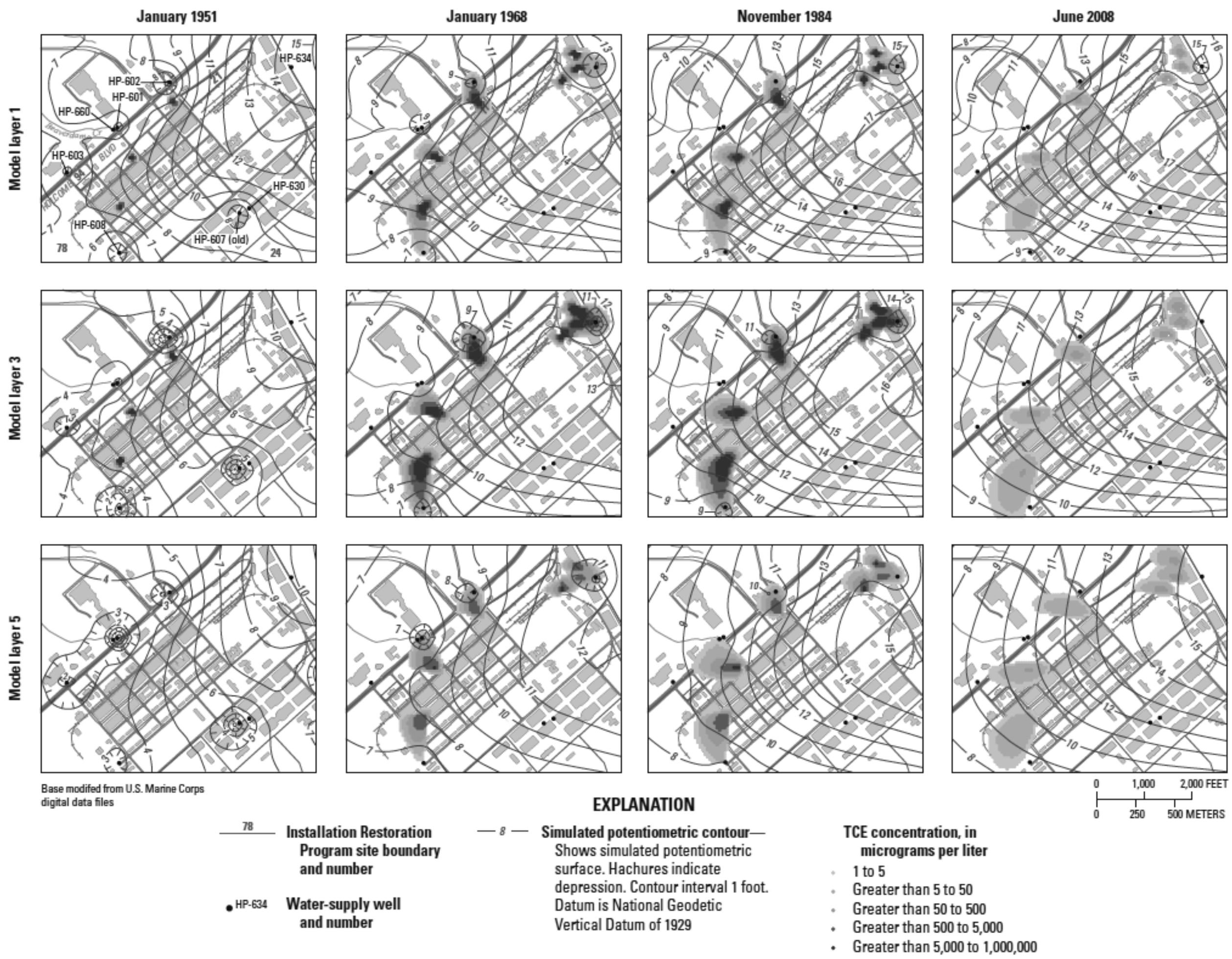

Figure A19. Reconstructed (simulated) water levels and distribution of trichloroethylene (TCE) within the Hadnot Point Industrial Area fate and transport model subdomain, model layers 1, 3, and 5, Hadnot Point–Holcomb Boulevard study area, U.S. Marine Corps Base Camp Lejeune, North Carolina, January 1951, January 1968, November 1984, and June 2008. (See Figure A13 for location and building numbers; see Appendix A4 for more detailed maps and results.)

On p. 81, Dr. Spiliotopoulos presents his Fig. 31 plotting of TCE concentrations in HP-634. However, he purposely does not include the data point with the value of 1,300 in his plot; including it would yield a very different picture, and show a much better match between simulated and observed TCE at the well location. TCE is found to be present in many locations immediately adjacent to HP-634, as seen in Fig. C33 (reproduced below). HP-634 is within the industrial area HPIA in that map (close to its northeastern boundary).

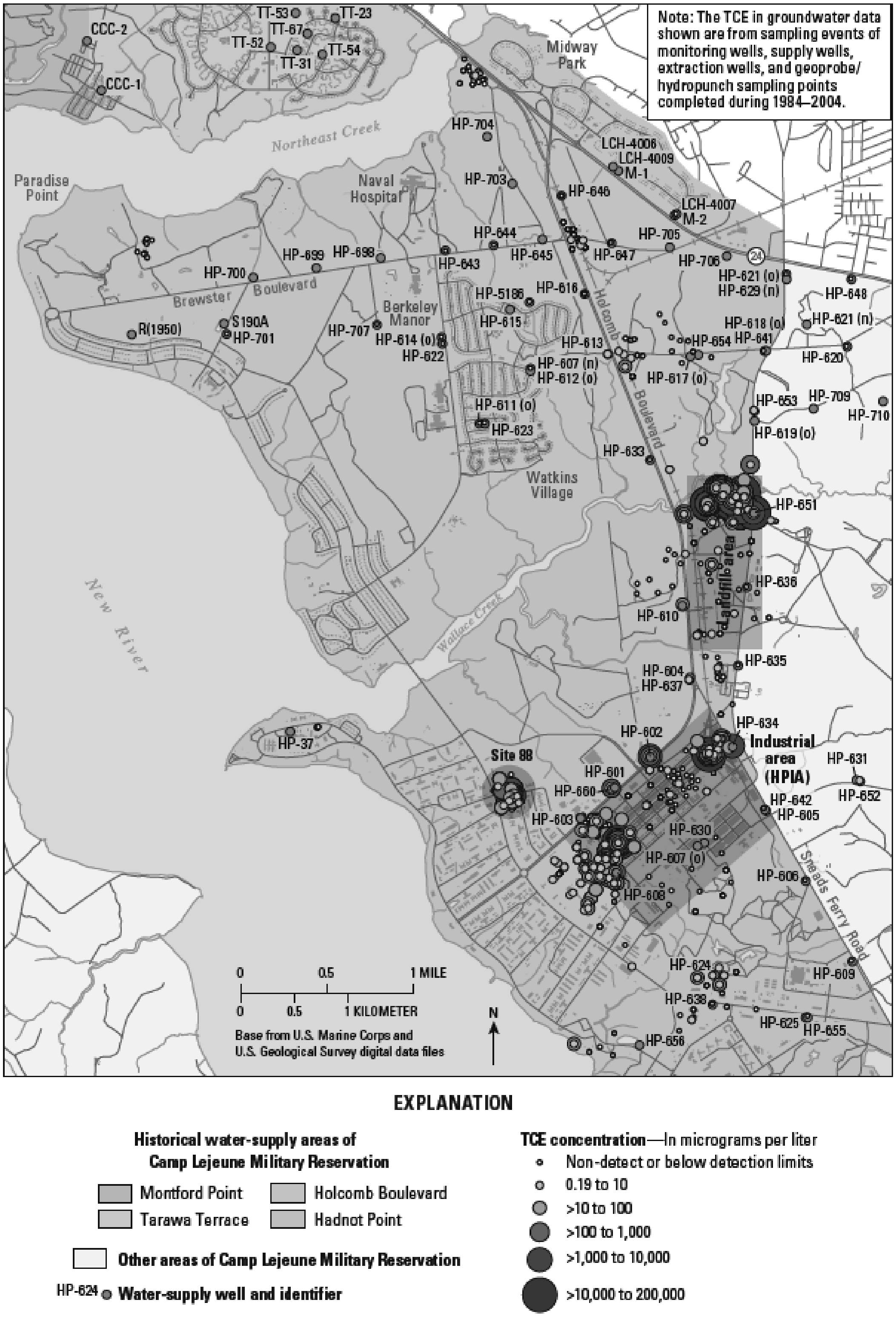

Figure C33. Groundwater sample locations for trichloroethylene (TCE) and ranges of TCE concentration in monitor and supply wells within the Hadnot Point–Holcomb Boulevard study area, U.S. Marine Corps Base Camp Lejeune, North Carolina.

Opinion 16: Dr. Spiliotopoulos argues here that the model for VOC degradation products was based on limited data, and "ATSDR's historical reconstruction prior to December 1984 cannot be verified." He cites section 4.2.4 as support.

In section 4.2.4 (p. 82-83), Dr. Spiliotopoulos states that "As illustrated in Figure 33 [ATSDR Fig. A25], the historical reconstruction prior to 1985 cannot be verified, due to lack of observed data for the period." This is true, and it is the reason why a simulation model was needed and was developed. For the four contaminants shown in Fig. 33, the agreement between simulated values and observed data is excellent in all four plots. This close agreement when observations are available builds confidence in the reliability of the model and its predictions, including for the hindcasting results for times prior to 1985.

In the summary for Opinion 16 (p. 83), Dr. Spiliotopoulos repeats that "… such data were not available prior to December 1984. Therefore, the estimated monthly contaminant concentrations cannot be verified." Again, the whole point was to use a technically sound model, which would be calibrated to available data in and after 1985, to estimate the values during the 15 or so years prior to that calibration period to inform the epidemiological studies. For PCE and TCE, the fit with the LCM model was actually slightly better than with the MT3DMS model, which was not designed to simulate degradation products. The quality of that fit is illustrated in Figure A25.

Opinion 17: Dr. Spiliotopoulos says that "the sensitivity analysis for the various contaminant sources in Hadnot Point indicated that the timing of source-release start date is uncertain and, therefore, it is impossible to determine the historical period that contamination was present in groundwater." The conclusion of this sentence does not follow from the precedent. Of course there is uncertainty in the timing of the release. That is well known. But the uncertainty does not make analyses impossible. Also, the uncertainty is not unconstrained. The model helps constrain the reconstructed history as it incorporates the physics of groundwater flow and solute (contaminant) transport. It is *not* impossible "to determine the historical period that contamination was present in groundwater." It can be (and was) estimated, but with the recognition of uncertainty in the model and in the predictions. There are a fair amount of data on the groundwater flow field, which provide the calibration basis for the flow model, and the calibrated flow model has sufficient accuracy and reliability to estimate groundwater velocities and directions. The model basically shows that to simulate the observed increases in concentration at observation points, the timing of the source release becomes more narrowly constrained and its uncertainty is reduced (but not eliminated). The key is that the flow model simulates groundwater flowpaths and velocities with reasonable and acceptable accuracy.

On p. 84, referring to underground storage tanks, Dr. Spiliotopoulos says "The empirical data for UST releases may or may not be applicable to the USTs installed at Camp Lejeune and, therefore, assignment of timing and magnitude for these sources is arbitrary and uncertain." Although uncertainty is clearly recognized, the assignment is not arbitrary. The basis is the EPA data on more than 12,000 leak incidents. Without direct observation to the contrary, why would one think that these USTs would behave much differently than the average failure time for such a large representative sample of documented cases? The approach used is not arbitrary, nor "highly" uncertain, nor an unreasonable assumption.

On p. 85, Dr. Spiliotopoulos goes on to discuss the range of years used in the sensitivity analysis, which spanned ±9 years. The point is not that the starting release date could have been anytime in that 18-year span, but rather to examine how sensitive the results are to such uncertainty. The results shown in Fig. 34 (ATSDR Fig. A37) indicate that at the later times—i.e., during the 18 years of the epidemiological studies—uncertainty in the starting release dates has little effect on estimated TCE concentrations. For the period between about 1950 and 1970, results from each of the various starting dates tend to converge on the same solution after only 3 or 4 years of simulation time.

In the summary for opinion 17 (p. 86), Dr. Spiliotopoulos says "it is not possible to confidently determine the actual period of groundwater contamination …" I would counter that it is possible to do so with some reasonable level of confidence, and ATSDR has done so. Of course there is uncertainty.

Opinion 18: Dr. Spiliotopoulos states that "the sensitivity analysis of the dose reconstruction model for HP was based on parameter variability unsupported by data." And that "the results of the sensitivity analysis were incorrectly presented as an uncertainty analysis range." Support is said to be in Section 4.2.5.1.2.

First, I note that there is some overlap and linkage between sensitivity analysis and uncertainty analysis. Anderson & Woessner (in their 1992 book on "Applied Groundwater Modeling") in discussing sensitivity analysis state: "The purpose of a sensitivity analysis is to quantify the uncertainty in the calibrated model caused by uncertainty in the estimates of aquifer parameters, stresses, and boundary conditions."

On p. 87 (Section 4.2.5.1.2) Dr. Spiliotopoulos argues that the sensitivity analysis used extreme values for parameters. But these "extreme" values were not used for the hindcasting (historical reconstruction), which was done using the calibrated model and calibrated parameter values. The wide range in parameter values was only used to assess model sensitivity and uncertainty, and thereby gain some further understanding of how and why the model is behaving as it does. This is not unusual. It has minimal or negligible effect on the calibrated model.

On p. 89, Dr. Spiliotopoulos argues that the range of parameter values in the sensitivity analysis was too wide. The inference then seems to be that the range of results (shaded areas) shown in Fig. 35 (ATSDR's Fig. A34, shown on p. 90) is too wide and should be narrower (closer to the results for the calibrated model). This doesn't seem like a major problem, as it would imply that the model results may be better defined than indicated otherwise. In looking at sensitivity, ATSDR did not imply that these "extreme" values were realistic or expected. They only illustrated a possible maximum bracketing of results.

In the Summary comments for Opinion 18, Dr. Spiliotopoulos concludes that "ATSDR presented the results of this analysis as indicative of the expected range of reconstructed monthly contaminant concentrations." I don't see where they said or implied this.

Opinion 19: Dr. Spiliotopoulos expresses a concern that the Hadnot Point analysis "only partially addressed model uncertainty." Support is included in Section 4.2.5.2.

In Section 4.2.5.2 (p. 91): In the first paragraph Dr. Spiliotopoulos seems to imply that ATSDR's use of Latin Hypercube Sampling was somehow an oversimplified approach. This is a valid and appropriate

method to use in these circumstances. For example, in conducting the Performance Assessment for the radioactive waste repository at the WIPP site in New Mexico, DOE and Sandia National Labs used the LHS approach with their groundwater flow and transport models for the WIPP site, as part of their application for approval to begin operations. This work was carefully reviewed by a National Academy Committee (NRC, 1996) and WIPP was granted approval to begin operations by the U.S. Environmental Protection Agency in the mid-1990s. There is nothing wrong (and a lot right) with the use of this method. EPA approval was granted even though there were no observations at all of concentrations in the aquifer of concern, yet predictions were made for 10,000 years into the future.

Section 4.2.5.2 (p. 91): In indicating that the uncertainty analysis was incomplete, Dr. Spiliotopoulos says (para. 2, p. 91) "ATSDR considered a small number of only 10 uncertainty scenarios." While it is debatable as to whether ten is a "small" number of scenarios to evaluate, it is a reasonable number to consider, and the 10 scenarios encompass a lot of the uncertainty in parameters and boundary conditions. ATSDR accomplished the goal of completing and documenting an uncertainty analysis, although it would have been possible to add additional scenarios to consider. It is highly unlikely, however, that adding more scenarios would lead to a modification of the calibrated model or to a different historical reconstruction.

In the first paragraph on p. 92, Dr. Spiliotopoulos quotes Doherty: "ideally, the value of the prediction should lie somewhere near the center of the uncertainty band." He then states that the ATSDR calibrated model "fails to conform with this rule …" However, this is not anyone's "rule." It is an idealization. Where the calibrated model lies off the center of the uncertainty range of estimates, it may simply be because additional parameters and scenarios need to be incorporated into the Monte Carlo simulations. In statistical testing, it is generally acceptable for a point or sample to fall within a range of two standard deviations of the mean.

In his summary for Opinion 19 (p. 92), Dr. Spiliotopoulos states that "the analysis only partially addressed the model uncertainty." But if more scenarios were considered or if more than 95% of the results were shown, the increased number of scenarios would widen the range and place the calibration results more consistently towards the middle of the range. Most of the time, the calibration is within the range of uncertainty brackets; when not, it is only very slightly above them. Overall, this does not seem to be a major issue. If additional factors were considered, the range would likely be wider and encompass all of the calibrated results. I also see no reason why this would have led to a different set of calibrated parameters.

Section 4.2.5.3, Concluding Remarks (p. 92): Dr. Spiliotopoulos reiterates his concern that there is lack of historical data to constrain the calibration. He quotes an article that says the "model should replicate observed system behavior." This must be taken in a general way because a model is by definition a simplified approximation of a complex real system, and no model can literally replicate a real system and its behavior. He argues that "The ATSDR model results did not meet this requirement." I disagree, and believe that there was a satisfactory representation of observed behavior for both head distributions and concentration distributions. Could it have been better? Sure, if more data had been available. Is it good enough to produce a reasonable hindcast historical reconstruction? I believe the answer is yes. Dr. Spiliotopoulos says "that there is 'no observed system behavior.'" This is simply wrong. There are some water-level data available, and very good agreement between observed and simulated heads (water

levels). This agreement provides confidence in the computed directions and velocities of contaminant migration. There are some observed concentrations. It would be nice if more concentration observations had been made in the past, but they weren't. Where such data are available, the model often provides a very good match to those data. With the goal and implementation of computing monthly averages, there is no way that the model could have replicated the large concentration changes sometimes observed over short time periods and between successive samples. He also states that "ATSDR failed to quantify the uncertainty range reliably." But they did quantify it and document it. They did so reliably. Perhaps it could have been more comprehensive and considered more factors, but that doesn't mean that they didn't "quantify it reliably." Although comprehensive uncertainty analysis is desirable, doing so is not a necessary condition for calibrating a groundwater model.

Section 4.2.5.3, Concluding Remarks (p. 93): Dr. Spiliotopoulos says "If parameter sensitivity and uncertainty can only be evaluated in a qualitative way, …" then the results and conclusions are not "scientifically defensible." The sensitivity and uncertainty analyses were definitely quantitative, and the quote from ATSDR (bottom p. 92) did not say these analyses were ONLY "qualitative". I believe that the model development by ATSDR for both TT and HPHB are scientifically defensible.

Review Comments on Chapter 3:

p. 10, Section 3.1.8 (Concluding Remarks): Dr. Spiliotopoulos says "Model calibration is not possible when there are no historical data to match." However, there are historical data available for Camp Lejeune. The ATSDR models were calibrated using comparisons to historical data—both groundwater level observations and some data on solute concentrations in water samples. There are many direct measurements of hydraulic conductivity—a key parameter in simulating groundwater flow and velocity. So the concluding statement above is simply not applicable to the ATSDR model development and calibration.

p. 12, Section 3.2: In this paragraph, Dr. Spiliotopoulos concludes by stating "However, the timing and quantification of contaminant releases from that source [ABC Cleaners] are uncertain, due to a lack of historical data." Of course, the timing and quantification of contaminant releases from ABC Cleaners has some associated uncertainty. However, there is knowledge of when they operated, precise information on its location, and there is little doubt that it was a source of contamination. The modeling exercises help reduce the uncertainty about the timing and strength of the contaminant source. It is rare (if ever) that the precise release dates and strengths of a historical contamination source are known. This is a type of uncertainty that is commonly dealt with in model development, and this type of uncertainty does not preclude the development, calibration, and usefulness of a groundwater model.

A related issue of contaminant travel times from ABC Cleaners to well TT-26: (**Hennet's report, p. 5-15 – 5-16 and his Attachment D**): Dr. Hennet estimates a range of values for travel times of PCE between ABC Cleaners and TT-26 that are stated to be "in the 15 to 25 years range", based on three assumed "representative" flow paths, indicating the arrival didn't occur until the 1970s. He presents supporting material and calculations in his Attachment D. Dr. Hennet assumes the horizontal travel distance in the shallow aquifer is either (1) 200 ft in the shallow aquifer and 800 ft in the pumped aquifer, (2) 500 ft in the shallow aquifer and 500 ft in the pumped aquifer, or (3) 800 ft in the shallow aquifer and 200 ft in the pumped aquifer. He further assumes that the hydraulic gradient in the layer 2 confining unit is the same

in all cases (i.e., at three different distances from the pumping well). This is not a reasonable assumption (for example, see TT Figs. C19 & C21). In the pumped aquifer, a cone of depression will form with lowest heads adjacent to the well and higher heads further from the well. In the shallow aquifer, the heads will not change much due to pumping in the deeper aquifer. This drawdown effect is strongest near the well, and results in a greater hydraulic gradient (and faster velocity) across the confining layer closer to the well.

Pumping also results in a steeper horizontal gradient (and faster velocity) closer to the well in model layer 3, and a shallower gradient further from the well. Dr. Hennet's calculations assume the same horizontal velocity in the pumped aquifer regardless of the distance from the pumped well, which is not a valid assumption.

Examining the heads for model layers 1 and 3 as shown in TT Figs. C18 and C19, and looking at a point about halfway between ABC Cleaners and TT-26 and at a point very close to TT-26, the head difference between the two layers (across the confining bed) is about 10' – 9' = 1 ft at the halfway location and about 5' – 2' = 3 ft at a location close to TT-26. Therefore, the hydraulic gradient potentially driving downward flow is about 3 times greater close to the well than it is halfway between the well and the contaminant source. So this large spatial change in vertical hydraulic gradient must be accounted for, and the assumption that it is the same at all locations cannot be supported. Dr. Hennet does not account for the steeper vertical gradient in layer 2 for the path closer to the pumped well, nor does he account for the faster velocity in layer 3 when the travel distance is only 200 ft.

It is more likely that the travel distance in the shallower aquifer for much of the contaminated shallow groundwater would be more than 800 ft and the corresponding travel distance in the pumped aquifer would be less than 200 ft because (1) the vertically downward transport is more likely to occur where the vertical gradient is the strongest in the confining layer, which is closest to the pumping well, (2) the downward velocity would be fastest where the gradient is steeper close to TT-26, and (3) according to Dr. Hennet's calculations, the downward flux is only about 5% of the horizontal flux in the shallow aquifer, so that even if some contaminant leaked downward at further upgradient distances from TT-26, much would remain in the shallow aquifer to migrate to locations closer to, or even adjacent to, TT-26, where downward leakage would be the fastest. Thus, Dr. Hennet's three "representative" flow paths did not include a more critical flow path in which travel in the shallower aquifer is close to 1,000 ft. For this critical flow path, the travel time would be much less than 15 years—on the order of 3.5 to 5 years. For these several reasons, Dr. Hennet's estimates of travel times from ABC to TT-26 are erroneous, misleading, biased-high, and based on unreliable assumptions.

Well TT-26 pumpage (Hennet's report p. 5-36): Dr. Hennet continues in criticizing the pumpage assumptions about well TT-26. He says, "ATSDR assumed that supply well TT-26 was constantly pumping prior to 1980. This is unlikely as supply wells cannot remain in service for decades without shut down periods for repairs and maintenance." Dr. Hennet implies it is unreasonable to assume this, yet offers absolutely no evidence to support his contention. This can be contrasted with ATSDR's study, which (p. 18) states that they have documented pumping records for TT-26 (and other wells) for some time periods and those estimates "are based on documented information detailing periods of maintenance for specific wells." For earlier periods in which there are no explicit pumping records, TT Chapter C (p. C22-C23) describes their estimation approach in detail (and Dr. Hennet does not offer a better way that this could have been done). Furthermore, in general, well maintenance frequently only requires a day to

a few days to complete. If TT-26 had been shut down for only a few days during a few months of every year for servicing, the monthly simulation model would still have to assume it operated for a full month each time, though at a proportionately reduced monthly pumping rate to reflect the actual total monthly withdrawal. It is hard to accept Dr. Hennet's speculative and hypothetical criticism or expect that it would make any difference.

p. 21-22 (Section 3.3) & p. 29: Dr. Spiliotopoulos cites Clement's 2011 issue paper (published in Ground Water journal); but these comments don't cite the Author's Reply (by Clement) to the published Comment by Maslia et al. in response to the original article. In his Reply to the Comment, Clement states "The goal of my article was not to review the Camp Lejeune (CLJ) modeling studies. Rather it was to use the CLJ problem as an example to highlight issues related to model complexities and to spark an open debate on when, where, and why we should limit model complexity." Therefore, Clement admits the article did not constitute a detailed technical review of the Camp Lejeune model study, so his 2011 Issue Paper that appeared to criticize it should not be taken as an expert analysis of the model or of its reliability or of the site. The Comment by Maslia et al. provided detailed rebuttals to Clement's concerns.

p. 21 (Section 3.3): Also, on p. 21 Dr. Spiliotopoulos states that "Dr. Clement's article echoed the NRC's concerns about the uncertainty in ATSDR's water model related to Tarawa Terrace and recommended a simpler approach for the water model related to Hadnot Point and Holcomb Boulevard to meet policy-oriented goals." Dr. Spiliotopoulos implies that the NRC report is a second independent review of the work. With regards to the groundwater modeling, it is not. Dr. Clement, a civil engineer, was the only groundwater expert on that committee (there were no geologists or hydrogeologists on that NRC Committee), so his concerns don't simply echo those of the NRC committee. Instead, it was likely that he was the source of those comments in the NRC Committee. While the use of "simpler models" might be okay for assessing policy-oriented goals, the simpler models would be subject to even greater uncertainty and lack of physical realism. Furthermore, the goals of historical reconstruction require a detailed and fairly complex modeling approach because the system being modeled is complex, and the use of simple models to meet such technical goals would be neither acceptable nor sufficiently accurate.

Regarding the 2009 NRC report and committee, Dr. Spiliotopoulos states that its primary charge was "to assess the strength of evidence in establishing a link or association between exposure to TCE, PCE, and other drinking-water contaminants and each adverse health effect suspected to be associated with such exposure." Consequently, almost all of the NRC Committee members were experts in medical and health fields. Only one was an expert in groundwater. The Committee had neither the focus, goal, intent, nor multiple experts to assess in depth the ATSDR's groundwater models. They were expected to focus on health effects.

Section 3.3 and scientific validity of ATSDR's models: In this section, Dr. Spiliotopoulos refers to statements by Dr. Dan Waddill. Dr. Waddill testified (Aug. 26, 2024, p. 234-235) regarding the ATSDR water modeling that "I do not think their results … were scientifically valid because, you know, science needs to be based on real-world observations and analysis. … and there were just not enough real-world measurements for this to count as a scientifically valid approach." He continues and concludes that the work was not scientifically valid because no concentration data were available in the 1950s-70s, and such observations can no longer be made (obviously). He argues that because of this, the hypothesis cannot be tested, so therefore it is not scientifically valid. I disagree.

I first note that Copi (1961) in discussing science and hypotheses states that "Few propositions in science are *directly* verifiable as true." He later states, "They can, however, be tested *indirectly*." Therefore, I would counter Dr. Waddill's statements by noting that in developing and applying the ATSDR groundwater models, that scientifically valid methods were used, and the models were based on sound hydraulic and physical principles that themselves have been tested and shown to be accurate and reliable approaches to describing and predicting groundwater flow and contaminant transport. The models were also based on many available hydraulic tests measuring hydraulic properties of the subsurface that do not change over time, and hence were data applicable to the site during the 1950s through 1970s. The models are indirectly tested during the calibration process in that available observations are compared to simulated values. This is an indirect type of model testing (or hypothesis testing) in which observations are compared to simulated values. The underlying theories and models have been tested in numerous field studies and are widely recognized as being scientifically valid.

The question should be whether this model for this site was sufficiently well calibrated and representative to perform a hindcasting prediction. I believe it was. I think there are many questions in our universe that are addressed using principles and models of physics that cannot (for all practical purposes) be directly tested in the foreseeable future. That does not render that work to be unscientific or lacking scientific validity. Predictive uses of models, whether forward in time or backwards in time, are widely accepted uses of scientifically valid models, while allowing for the existence and recognition of uncertainty in those predictions. The fact that there is uncertainty does not mean that they are not scientifically valid or scientifically defensible. The fact that one type or time period of observations are not available does not mean that the model is not scientifically valid.

Section 3.3 and Calibration Targets: At several places in this section, the issue of "calibration targets" is mentioned along with criticism that some simulated values did not fall within the calibration target. Relevant to this discussion are my comments in the 2009 Expert Panel Report (p. 101), with which I still agree and which I therefore repeat verbatim here:

> "a. **Are there established standards for establishing specific calibration targets? If so, what are they?** Overall, there are no standards and probably should not be any. Such targets are inevitably arbitrary and to some extent meaningless. They tend to distract from the quality of the calibration process and shift focus to the arbitrary goal. It is a "red herring." Not achieving a predetermined calibration target should not disqualify a model, nor does that prove a model is not valuable or useful. Conversely, meeting such a predetermined calibration target does not prove that the model is a good one or that it meets the needs of the particular study or that its calculations and predictions are accurate and/or reliable.
>
> "b. **Should ATSDR establish different calibration targets than for the Tarawa Terrace model?** In my opinion, the use of specific calibration targets should be abandoned. They have no real value in the context of hydrogeology, and can only serve to provide a false or meaningless image of the quality of the developed model. ATSDR only has a limited time to complete the study, and you will do the best job possible within that limited time and budget. Applying a calibration target will not lead to a better model, but it will cause some time to be spent on comparing the results to the target, and perhaps forcing the results to

fall within the target. It would be better to include on-going independent expert peer review during the model development process, as this will have a much higher payoff than calibration targets in terms of improving the quality of the final product."

Conclusions:

Groundwater models must be (and have been) calibrated in the absence of early time concentration data, as ATSDR has done. Other representative published examples where this has been successfully accomplished include the Rocky Mountain Arsenal, CO (Konikow, 1977) and Lawrence Livermore National Laboratory, CA (Rogers, 1992). In both of these cases, the early time history was reconstructed as part of the model calibration process (it just wasn't called "hindcasting"). This is a widely accepted procedure among groundwater modelers.

Although Dr. Spiliotopoulos repeatedly questions the accuracy of the ATSDR model and its calibration, I don't see any evidence that it is unacceptably inaccurate. In my opinion, ATSDR followed generally accepted methods that yielded reasonably accurate results for the mean monthly concentration of contaminants. ATSDR's TT Table F13 shows comparisons between observed and simulated concentration values, and most (but not all) are within the calibration target range. The presence of differences is not unexpected and does not indicate the model is unreasonably inaccurate or unscientific. Concentrations for many chemical constituents in groundwater typically show a high variation at local spatial scales and small time scales—much greater variability than presented by hydraulic heads. This is normal, and no groundwater transport model would be expected to reproduce or explain such small-scale variability in concentration.

Dr. Hennet presents a summary opinion on p. 5-36 of his report stating "ATSDR's assumptions are deficient, not verifiable, and at times demonstratively incorrect." I believe, to the contrary, that ATSDR's assumptions are reasonable and clearly documented with their supporting basis clearly described in detail and with recognition of uncertainty. I would argue that his counter examples, such as for bulk density and K_d, make little to no difference. Dr. Hennet's own estimates of travel times are clearly deficient and incorrect. Of course, the early time reconstructed concentrations cannot be directly verified. Those data don't exist. That is why the state-of-the-art simulation models were needed. He further states that "ATSDR estimates are not quantitatively reliable as different plausible assumptions would lead to different results." Nonuniqueness of calibrated groundwater models is a well-recognized issue. Different assumptions can lead to different results and different assumptions can also lead to identical results. This is true of every groundwater model ever developed. It does not negate the value or reliability of the model. This is why sensitivity and uncertainty analyses are helpful. Furthermore, it is why we put strong reliance on the expert judgment of those who have studied the particular aquifer system the longest and most in-depth, such as the ATSDR's authors of the modeling reports. Finally, Dr. Hennet says "ATSDR COC concentration estimates are for raw water which is not equivalent to COC concentrations in the distributed water." As I previously stated above, the opinion of experts on the 2005 Expert Review panel was that possible COC losses during water treatment at the Camp Lejeune WTPs would be small to minimal.

In my opinion, ATSDR has done an admirable job in completing a challenging task of using hindcasting with a calibrated model to reconstruct credible concentration distributions in time and space prior to the

availability of data from chemical analyses of groundwater samples in the mid-1980s. In the face of missing historical data, the ATSDR models provide useful input to epidemiological studies. ATSDR clearly and comprehensively documented the model development—providing transparency to their work. There is uncertainty in the calibrated models (as there always is in such models) and in the hindcasted results, and that is clearly recognized and evaluated. The uncertainty is not so large or unexpected as to preclude the use of the model results in the epidemiological studies or for providing monthly mean concentrations for use by health professionals to estimate past exposure of residents on an "as likely as not" or "more likely than not" basis. The methods used were rigorous and scientifically sound.

Leonard F. Konikow

Dr. Leonard F. Konikow, PhD, NAE

January 13, 2025

References

Anderson, M.P. and Woessner, W.W., 1992, Applied Groundwater Modeling: Simulation of Flow and Advective Transport. Academic Press, San Diego. 381 pp.

Copi, I.M., 1961, Introduction to Logic. [2nd Ed.] Macmillan, New York. 512 pp.

Helsel, D. and Lee, L., 2006, Analysis of Environmental Data with Nondetects: Statistical Methods for Censored Environmental Data. Continuing Education Workshop at the Joint Statistical Meetings, American Statistical Association, Seattle, Washington (Aug. 2006). [Based on the textbook "Nondetects and Data Analysis: Statistics for censored environmental data" By Dennis Helsel, Wiley, 2005.]

Hill, M.C. and Tiedeman, C., 2007, Effective Groundwater Model Calibration, with Analysis of Data, Sensitivities, Predictions, and Uncertainty. Wiley-Interscience, Hoboken, NJ, 455 pp.

Konikow, L.F., 1977, Modeling chloride movement in the alluvial aquifer at the Rocky Mountain Arsenal, Colorado: U.S. Geol. Survey Water-Supply Paper 2044, 43 p.

Konikow, L. F., and Thompson, D.W., 1984, Groundwater contamination and aquifer reclamation at the Rocky Mountain Arsenal, Colorado: *in* Groundwater Contamination, National Academy Press, Washington, D.C., p. 93-103.

Kret, E., Kiecak, A., Malina, G., Nijenhuis, I., and Postawa, A., 2015. Identification of TCE and PCE sorption and biodegradation parameters in a sandy aquifer for fate and transport modelling: batch and column studies. Environ. Sci. Pollut. Res. 22. p. 9877–9888. DOI 10.1007/s11356-015-4156-9.

National Research Council. 1996. *The Waste Isolation Pilot Plant--A Potential Solution for the Disposal of Transuranic Waste*. National Academy Press, Washington, D.C., 169 p.

National Research Council. 2009. *Contaminated Water Supplies at Camp Lejeune: Assessing Potential Health Effects*. Washington, DC: The National Academies Press. ttps://doi.org/10.17226/12618.

Reilly, T.E., and Harbaugh, A.W., 2004, Guidelines for evaluating ground-water flow models: U.S. Geological Survey Scientific Investigations Report 2004-5038, 30 p.

Rogers, L. 1992. History matching to determine the retardation of PCE in ground water. *Groundwater* **30**, no. 1: p. 50–60.

Zheng, C., and Bennett, G.D. 2002, Applied Contaminant Transport Modeling. [2nd Edition] Wiley-Interscience, New York. 621 pp.

ATTACHMENT A

Leonard F. Konikow, PhD, NAE

Contact Information

11316 Myrtle Ln, Reston, VA 20191; 703-855-5498 (Cell); email: LFKonikow@gmail.com

Education

1966 BA, Geology, Hofstra University, Hempstead, New York
1969 MS, Geology, Pennsylvania State University
1973 PhD, Geology, Pennsylvania State University

Work Experience

2020-2023 Editor-in-Chief, *Groundwater* journal
2013-2023 Emeritus Scientist, Water Mission Area, U.S. Geological Survey
1980–2013 Project Chief, Natl. Research Program, Water Resources Division, U.S. Geological Survey, Research Project "Digital modeling of transport in saturated zone"
1978–1980 Staff Hydrologist, Ground Water Branch, U.S. Geological Survey, Reston, Virginia
1974–1978 Project Chief, Research Project "Solute Transport in Ground Water," U.S. Geological Survey, Central Region, Lakewood, Colorado
1972–1974 Project Chief, Subsurface Waste Investigations, U.S. Geological Survey, Lakewood, Colo.
5/66-7/66 Hydrogeologist assistant, Geraghty & Miller, Inc., Groundwater Consultants, Port Washington, NY.

Other Experience*

Instructor and lecturer at:
Fall 1991 & 1992 Department of Environmental Sciences, University of Virginia
Fall 1997 Department of Geological Sciences, Stanford University

Professional Societies*

- American Geophysical Union (AGU) (1970–present; elected Fellow, 2001)
- AGU Spring Meeting Program Chairman for Hydrology (1984–1987)
- AGU Groundwater Committee (1977–1986; Chairman, 1980–1982)
- Geological Society of America (GSA) (1974–present; Fellow since 1990)
- Management Board, Hydrogeology Division, GSA (1991–1995); Chair (1993-1994)
- International Association of Hydrogeologists (IAH) (1985–present); IAH Vice President, North America and IAH Executive Council (2009-2012); Chairman of U.S. National Chapter, IAH (2001–2004)
- National Ground Water Association (NGWA), Scientists and Engineers Division (SAE) (1990–present)
- NGWA/SAE—Board of Directors (1996–2000)
- American Institute of Hydrology (Certified Professional Hydrogeologist) (1991–2023)
- Registered Professional Geologist, Pennsylvania (1996–2023)

Honors and Awards*

- Birdsall Distinguished Lecturer (1985–1986), GSA, Hydrogeology Division
- M. King Hubbert Science Award (1989), Life Member Award (2013) National Ground Water Association
- O.E. Meinzer Award (1997), Distinguished Service Award (2000), GSA, Hydrogeology Division
- C.V. Theis Award (1998), American Institute of Hydrology
- Distinguished Service Award (1999), U.S. Department of Interior
- Elected as Fellow (2001), American Geophysical Union
- President's Award (2001), International Association of Hydrogeologists
- **Elected to National Academy of Engineering (2015)**

*** Partial list; more details available upon request.**

Selected Professional Activities*

- Rocky Mountain Arsenal (Colorado) Technical Review Committee (1975–1977)
- National Research Council, Panel on Groundwater Contamination (1981–1982)
- National Research Council, Committee on Ground-Water Modeling Assessment (1987–1989)
- National Research Council, Waste Isolation Pilot Plant Committee (1989–1997)
- Peer Review Panel, U.S. EPA Environmental Monitoring Systems Lab, Las Vegas, Nevada (1991)
- National Science Foundation, Review Panel for Hydrologic Sciences and interim staff assistant (1992)
- Member of Modeling Project Subcommittee, Science Advisory Board, U.S. EPA (1993)
- Associate Editor, Ground Water Journal (1993–1995); Water Resources Research (1981-1984)
- Adviser to U.S. AID project studying seawater intrusion in Gaza and Morocco (1994–1997)
- National Research Council, Committee on Principles and Operational Strategies for Staged Repository Systems (2001–2002)
- Expert Peer Review Panel for ATSDR to evaluate historical ground-water contamination and water-supply distribution problems at Tarawa Terrace, U.S. Marine Corps Base, Camp Lejeune, NC (March 2005)
- Expert Peer Review Panel for the South Florida Water Management District (SFWMD) to evaluate East Central Florida Transient model (Oct. 2006-Feb. 2007)
- Coastal Sound Science Initiative Technical Advisory Committee for Georgia and S.C. (2008)
- Expert Peer Review Panel for ATSDR to evaluate historical ground-water contamination at Hadnot Point area of the U.S. Marine Corps Base, Camp Lejeune, NC (April, 2009)
- International Scientific Advisory Committee for the National Centre for Groundwater Research & Training of Australia (2009-2014)
- Member of an Independent Review Panel to evaluate the Death Valley Regional Flow System Project for the USGS Nevada Water Science Center (March-July 2014)
- AGI, Geoscience Policy Advisory Comm. (2005-16); Critical Needs Working Group (2015-16)
- Invited lecturer, McCormick Specialized Reporting Institute 2015: Covering Water in a Changing World, University of Florida, College of Journalism & Communications, Gainesville, FL (Nov. 2015)
- Chair, DOE Expert Peer Review Panel for the Rainier Mesa/Shoshone Mountain Flow and Transport Model, Nevada National Security Site (NNSS), Nevada (Oct. 2017-April 2018)
- Member, DOE Expert Peer Review Panel for the Pahute Mesa Flow and Transport Model, NNSS (2022)

Publications*

Author or coauthor of numerous articles in peer-reviewed journals, government publications, conference proceedings, book chapters, and talks given at professional society meetings (complete detailed list available on request).

Konikow, L.F., and Bredehoeft, J.D., 1974, Modeling flow and chemical-quality changes in an irrigated stream-aquifer system: Water Resources Research, v. 10, no. 3, p. 546-562.

Konikow, L.F., and Bredehoeft, J.D., 1978, Computer model of two-dimensional solute transport and dispersion in ground water: U.S. Geol. Survey Techniques of Water-Resources Investigations, Book 7, Chapter C2, 90 p.

Konikow, L.F., 1986, Predictive accuracy of a ground-water model -- Lessons from a postaudit: Ground Water, v. 24, no. 2, p. 173-184.

Konikow, L.F., and Bredehoeft, J.D., 1992, Ground-water models cannot be validated: Adv. Water Res., v. 15, no. 1, p. 75-83.

Ewing, R.C., Tierney, M.S., Konikow, L.F., and Rechard, R.P., 1999, Performance assessments of nuclear waste repositories: A dialogue on their value and limitations: Risk Analysis, v. 19, no. 5, p. 933-958.

Konikow, L.F., and Kendy, E., 2005, Groundwater depletion: A global problem: Hydrogeol. Jour., v. 13, no. 1, p. 317-320.

Konikow, L.F., 2011, The Secret to Successful Solute-Transport Modeling: Ground Water, v. 49, no. 2, p. 144-159. doi: 10.1111/j.1745-6584.2010.00764.x.

Bredehoeft, J.D., and Konikow, L.F., 2012, Reflections on Our Model Validation Editorial: Ground Water, v. 50, no. 4, p. 495.

Konikow, L.F., and Leake, S.A., 2014, Depletion and Capture: Revisiting "The Source of Water Derived from Wells": Groundwater, v. 52, no. S1, p. 100-111. DOI: 10.1111/gwat.12204.

Konikow, L.F., 2015, Long-term groundwater depletion in the United States: Groundwater, v. 53, no. 1, p. 2-9.

Konikow, L.F. and J.D. Bredehoeft, 2020. Groundwater Resource Development: Effects and Sustainability. The Groundwater Project, Guelph, Ontario, Canada. https://gw-project.org/books/groundwater-resource-development/

* **Partial list; more details available upon request.**

ATTACHMENT B

Jan. 13, 2025

Leonard F. Konikow

Publications During Past 10 Years

Konikow, L.F., 2015, Long-term groundwater depletion in the United States: Groundwater, v. 53, no. 1, p. 2-9, doi: 10.1111/gwat.12306.

Alley, W.M., and Konikow, L.F., 2015, Bringing GRACE Down to Earth: Groundwater, v. 53, no. 6, p. 826-829, doi: 10.1111/gwat.12379.

Simmons, C.T., Anderson, M.P., Bredehoeft, J.D., Freeze, R.A., Gorelick, S.M., Konikow, L.F., de Marsily, G., Sudicky, E.A., and Voss, C.I., 2015. *Call for an International Strategic Plan for Groundwater*, Aqua 2015 - Hydrogeology: Back to the Future, 42nd International Congress of the International Association of Hydrogeologists, September 13-18, Rome, Italy.

Konikow, L.F., 2020, Passing the Torch. [Editorial] Groundwater, v. 58, no. 1, p. 1-2. doi: 10.1111/gwat.12964.

Marshall, S.K., Cook, P.G., Konikow, L.F., Simmons, C.T., and Dogramaci, S., 2020, Conjoint use of hydraulic head and groundwater age data to detect hydrogeologic barriers. *Hydrogeology Journal,* v. 28, 1003–1019. https://doi.org/10.1007/s10040-019-02095-9

Konikow, L.F. and J.D. Bredehoeft, 2020. Groundwater Resource Development: Effects and Sustainability. The Groundwater Project, Guelph, Ontario, Canada. https://gw-project.org/books/groundwater-resource-development/.

Konikow, L.F., 2021, Science and Politics. [Editorial] Groundwater, v. 59, no. 3, p. 310–311. doi: 10.1111/gwat.13097.

S.K. Marshall, P.G. Cook, C.T. Simmons, L.F. Konikow, and S. Dogramaci. 2022. An Approach to Include Hydrogeologic Barriers With Unknown Geometric Properties in Groundwater Model Inversions. Water Resources Research, 58(7). https://doi.org/10.1029/2021WR031458

Scanlon, B.R., Fakhreddine, S., Rateb, A., I. de Graaf, J. Famiglietti, T. Gleeson, R.Q. Grafton, E. Jobbagy, S. Kebede, S. Kolusu, L.F. Konikow, *et al.* 2023. Global water resources and the role of groundwater in a resilient water future. *Nat. Rev. Earth Environ.* **4**, 87–101. https://doi.org/10.1038/s43017-022-00378-6

L.F. Konikow, E.G. Reichard, and M.J. King, 2023, John D. Bredehoeft (1933-2023)—A Geologist at the Frontiers of Hydrogeology. Groundwater, v. 61, no. 3, p. 446-451. https://doi.org/10.1111/gwat.13313.

Konikow, L.F., 2023, Peer Review: Again. [Editorial] Groundwater, v. 61, no. 3, p. 289-290. doi: 10.1111/gwat.13312.